UNIVERSAL LIBRARY

DEDICATION

To the Engineering Students of the University of California, past, present and future, and especially to the students in Physics 1C during the years 1924 to 1930 inclusive, for whose benefit and with whose cooperation this book was written, the book is appreciatively dedicated.

Berkeley, California,
December 10, 1930.

PREFACE

THE engineering curriculum of the University of California requires a two-year course in college physics (four semesters, each containing three units' work) of all engineering students. In the first two semesters (freshman year) the subjects studied are mechanics, properties of matter, and heat. In the first semester of the second year electricity and magnetism are taken up, and in the second semester sound and light are studied. At the beginning of the second year these students have already taken analytic geometry and differential calculus, and, concurrently with the course in electricity, they are taking up the integral calculus. The purpose of these courses is to give a fundamental working knowledge of the physical principles with relatively little material of a technical engineering nature. Each of these semester courses consists of two one-hour demonstration lectures and a three-hour laboratory period each week during the fifteen weeks of the semester. During the first year the first hour of the laboratory period is taken up by a conference in groups of twenty students under an instructor, in which the experiments to be performed are discussed, as well as the problem sets which were to have been solved during the week. In connection with the lecture portion of the course the students are required to work twelve weekly problem sets (consisting of three to five problems each) during the semester. The lecture portion of the course also includes two or three mid-term examinations, each lasting an hour, and a three-hour final examination at the end of the semester.

In 1924 the author was given the task of conducting one of the two sections of this course. In the ensuing year both sections, comprising, all told, some three hundred engineering students, were given him to instruct. The task of teaching this course presented a difficult problem for the following reasons. It is the opinion of the author, gained from his first teaching experiences under the system initiated by Professor R. A. Millikan at the University of Chicago, that in a course of college physics the interesting phenomenology taught in high school physics and in some colleges, should be replaced by a study of the quantitative basis of the subject. In such a formulation the quan-

titative laws of physics cannot be adequately discussed without the use of mathematics. To engineering students, who will have to apply these quantitative laws in practice, it seems especially essential that the mathematical approach to the subject should be emphasized and taught in such a way as to enable the student to apply his physics to the practical problems which he meets in his engineering. Thus, despite the modern tendency of college physics texts to shirk mathematical formulations for the more dramatic and easy qualitative treatment, the author was committed, in attempting to teach this course, to a presentation of the subject of electricity and magnetism in a more quantitative and mathematical fashion.

The problem raised by the presentation of this course in accord with the above principle was the following. The course consists of twenty-six one-hour demonstration lectures which must be presented in such a form that all the mathematics used is available to the student for application to the problems to be solved. It was further complicated by the fact that, in the author's opinion, no law of physics or quantitative treatment of a phenomenon should be dealt with on the basis of what appears to the student to be an empirical equation. Every equation and quantitative relationship appearing in a text, in the author's opinion, should be justified on the basis of the fundamental experiments proving this law, or where it is a derived expression, the derivation should be given. Thus it is obvious that it is impossible to perform the demonstration lecture experiments and explain them and at the same time place on the board, in a sufficiently concise and clear form for the student to include in his lecture notes, the derivation of the necessary laws. An adequate text or reference book was therefore urgently needed. A rather thorough survey of the college texts in physics existing showed the author that the books which adequately covered the material required of these students were either so old as to be completely out of touch with the modern developments of physics or else entirely too advanced for a course of the nature above outlined. Of the modern college texts, in view of the tendency to shirk the more mathematical formulation, very few were adequate, and the few which gave some mathematical formulation were not complete enough in fields in which instruction was required. The author was therefore forced to prepare for the benefit of his students a condensed set of lecture notes which would fulfill the demand in such a fashion that the major portion of the lecture period could be given over to the performance of the demonstrations and their explanation.

The text was dictated from rough notes during a brief period one summer. It was revised the following year, and the present volume

forms the fourth rather complete revision of the original material. In making these revisions the author wishes to acknowledge assistance and suggestions concerning the text from Professor F. K. Richtmyer of Cornell University, Professor J. C. Hubbard of Johns Hopkins University, from his colleague Professor Thomas Buck of the Department of Mathematics of the University of California, from his friend Professor David L. Webster of Stanford University, who was kind enough to glance through the sections of the text, and, finally, from his wife Lora Lane Loeb, who has assisted him in the various revisions and in the preparation for publication in book form. He particularly wishes to acknowledge his indebtedness to his former teacher, Professor E. E. Hall, Chairman of the Physics Department of the University of California, under whose direction the course was first developed and whose plan of attack was, in general, followed in these notes.

In presenting the text, possibly a few words might be said concerning the plan which is followed and some of the more important points which the author has wished to emphasize. The general plan of presentation of the subject and the sequence of development of the materials of the course were not entirely left to the choice of the author. The sequence was influenced by considerations of laboratory equipment and facilities which could not be changed. For this reason the approach to the subject of electricity is not made from the point of view of static electricity, as the author wishes it might be. Aside from this, however, the general development and method of approach is one that is not infrequently used and it seems to have worked out fairly successfully in practice.

Again, in developing the subject with the purpose of fixing in the minds of the students the fundamental elements of electricity and magnetism, the author has found that the greatest difficulty encountered by the students is the confusion in the definition and relation of the many electrical entities occurring in such a treatment. To systematize the teaching of these fundamental concepts the author has found it convenient to organize the course about a skeleton based on a clear definition of the electrical entities and their interrelations. In doing this he has found it convenient to divide these entities, where they appear in the development of the subject, into two types, *fundamental* and *derived*. The entities chosen as *fundamental* are defined directly in terms of concepts of force and work and are electrical current, quantity, and potential. These definitions are on the basis of fundamental experiments relating these concepts to force and work. The three derived entities which he has chosen result from *ratios of*

the fundamental entities and are generally properties of the shape, dimensions, and materials of the electrical system where they occur. They are resistance, capacity, and self-induction. These six entities are, for practical purposes of measurement and comparison, expressed in three systems of units, giving, wherever possible, the reason for the origin and choice of the units. These systems of units are the absolute electromagnetic, the practical and the absolute electrostatic systems of units. This formulation of the fundamental notions of electricity leads to a summary, which constitutes the kernel of the course, based on the fundamental electrical entities and their relations. These are contained in the apparently elaborate table at the end of the text. Practice has shown that this bird's-eye view of the interrelations of the electrical concepts has been exceedingly helpful to the students. It leads to a mnemonic system for remembering the units and their ratios which has many interesting features. One of the interesting results from this choice of definitions leads to the observation that the ratios of the *fundamental* electromagnetic and electrostatic units all involve the *first power of the velocity of light*, while the ratios in the electromagnetic and electrostatic systems of the *derived units*, as defined above, involve the *square of the velocity of light.*

The nature of the training required of the California engineer in this course requires a thorough knowledge of Kirchhoff's laws of divided circuits, as well as a fundamental knowledge of the laws of magnetic circuits. These are not generally presented in college texts in such a form as to be practically applicable to the solution of problems. A great deal of care has therefore been taken in working out a system of application with illustrations for the benefit of the students.

Another feature of this text is the introduction, wherever possible, of the modern atomic and electronic theory of matter in explaining the phenomena. This is particularly noticeable in the discussion of the laws of electrolysis in terms of the octette theory of valence and in magnetism. The author believes that the treatment of the electrical cell presented is a new mode of approach to the study of this problem for physics texts, though it has been used to some extent by the chemists.

In preparing the text for publication in book form, the author has made two important changes as a result of suggestions made by Professor Hubbard. He has very much enlarged and extended the historical chapter. This was done first because the present status of physical science requires, as a result of relativity, quantum theory

and wave mechanics, that we should, at this time, take stock of the advance of physical science and its aims, past and present. In the second place it was felt that the history of electricity and magnetism, especially after 1830, has been very much neglected. It has thus been the aim to present the historical survey in two parts, first a general discussion of the development of physical science and secondly a more detailed story of the advance of electricity and magnetism in which the subject is conveniently divided into seven periods.

At the suggestion of Professor Hubbard a new chapter on the physical basis of thermionics and of photoelectric phenomena has been added, the reason for this being that the modern engineering training requires a knowledge of these subjects in later work. In this presentation the fundamental knowledge is largely stressed, including the influence of the modern theory of the electronic state in metals. Applications have, as is consistent with the spirit of this text, been included only sufficiently to illustrate the physical principles.

The author desires to acknowledge the invaluable assistance which he received in writing the historical survey from the use of two books, viz., Edmond Hoppe's "Geschichte der Physik" and Park Benjamen's "History of Electricity and Magnetism up to 1750." In addition, the author desires to acknowledge indebtedness to the admirable text of Dr. H. J. Van der Bijl, entitled "The Thermionic Vacuum Tube," in preparing the second part of Chapter XXVII on certain phases of the application of thermionics.

Finally it seemed to the author that with the training which the students in this course are having in mathematics, the application of their knowledge to the study of the physical problems which they are taking up is not amiss. Thus the author has not hesitated to introduce the use of a few simple derivatives and integrals where needed, although in doing this he has been careful to explain the physical meaning of the operations. As a particular feature of this treatment the more complete introduction to transient phenomena and to the rôle of inductance and capacity in sinusoidally alternating currents has given an admirable opportunity. The applications of the mathematics here used have been carefully looked over by the professor in charge of the mathematical instruction of these same students and have his hearty endorsement.

LEONARD B. LOEB.

PHYSICAL LABORATORY,
UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA,
October, 1930.

CONTENTS

GLOSSARY OF NOTATION

Amperes of current	i
Amplifying factor, thermionic	μ
Angle	θ
Angle, phase	ϕ
Area	A
Area, element of	dA
Average represented by a dash over a letter, e.g	$\bar{v}$
Avogadro number	N_A
Boltzmann constant	k
Capacity	C, C_1, C_2
Constant of galvanometer	K or K'
Constant of gravitational acceleration	g
Constant of integration	k
Current, absolute units	i_a
Current, amperes	i
Damping factor	ρ
Density, surface, of electricity or magnetism	σ
Density, volume, of electricity	ρ
Dielectric constant	D
Dimensions: Mass	M
Length	L
Time	T
Distance (radial)	r, r_1, r_2
Distance perpendicular to radius	s
Distance (length)	l, d, h
Distance, element of	ds, dl, dr
Efficiency	Eff
Electric field	F, F_1, F_2, or else X
Electrical quantity, absolute units	q, q_1, q_2
Electrical quantity, coulombs	Q, Q_1, Q_2
Electrochemical equivalent	S
Electromotive force, EMF	E
Electromotive force (back)	E'
Electron, charge of	e or sometimes ϵ
Exponential function	e
External resistance	R_e
Field, electric	F, F_1, F_2
Field, magnetic	H
Figure of merit	k
Flux, magnetic	ϕ
Force	f, f_1, f_2, f_m, f_h
Force, resultant	f_r

Force moment, total	G
Force moment, partial	g, g_1, g_2, etc.
Frequency of electromagnetic radiation	ν
Impedance	z
Induction, magnetic	B
Inertia, moment of	I
Internal resistance	R_i
Intensity of magnetization	I
Logarithm, to the base e. (The only logarithms used in text)	log
Magnetic moment	M
Magnetic permeability	μ
Magnetic pole strength	m, m_1, m_2
Magnetic susceptibility	κ
Magnetomotive force	M.M.F.
Multiplying factor	X
Mutual inductance	M
Number, number of turns	n, n_1, n_2
Number of revolutions per second	N, N_1
Ohms of resistance	Ω
Quantum action, unit of	h
Peltier coefficient	p
Period	T
Period factor	p
Pi	π
Planck constant of quantum action	h
Potential difference, in volts	V, V_1, V_2
Potential difference, in absolute units	$P.D.$, $P.D._1$, $P.D._2$, $P.D._3$, etc.
Power	P
Quantity of electricity, absolute	q, q_1, q_2
Quantity of electricity, in coulombs	Q, Q_1, Q_2
Quantity of heat	H
Quantity, magnetic	m, m_1, m_2
Reluctance	Z
Resistance	R, R_1, R_2, R_3
Resistance at 0° C	R_o
Resistance specific	R_s
Self-induction	L
Temperature in degrees C. or absolute	T
Thermionic work function	χ or b
Time	t
Time interval	τ, dt
Torsional constant	T_o
Virtual volts, or amperes	V_v or i_v
Velocity of electrons or molecules	v
Velocity of light	c
Work	W

FUNDAMENTALS OF ELECTRICITY AND MAGNETISM

CHAPTER I

HISTORICAL

1. GENERAL DEVELOPMENT OF PHYSICAL SCIENCE

IN introducing the subject of electricity and magnetism it might not be amiss to do so with a brief historical sketch; for it often happens that such a sketch brings out the development of ideas in relation to each other and to the subject as a whole with a clarity which is generally lost sight of in the more detailed study of the individual parts. This bird's-eye view of the subject is often most stimulating, for it not only relates the development of electricity to the development of other branches of physics as a whole, but also shows how inevitably progress or discovery is linked up with and depends on antecedent progress or discovery.

Before the time of the Greeks, say before 600 B.C., there was no such thing as science in the present sense of the word, nor even in the earlier Greek meaning of the word. What body of knowledge existed consisted of certain arbitrary uncoordinated facts gleaned by independent discovery, invention, or observation. These facts existed as isolated items, unrelated to each other or to any body of knowledge. In the Rhind papyrus of 2000 B.C. we see, therefore, merely a compilation of isolated empirical facts.

Beginning with the Pythagorean school, and perhaps a few decades before, we see the first recorded attempts made to coordinate existing observation and knowledge into schemes of universal scope, i.e., philosophical systems. That is, mankind had at this point begun to assume vast causal relations of universal extent to exist in Nature and was attempting to discover or divine these. As an example, Pythagoras observed that the seven strings of the Greek lyre when equally taut emitted a harmonious system of notes when struck. Since Pythagoras believed that all things could be expressed by numbers, he assumed

that the ratios of the lengths of the strings of the lyre must represent a set of harmonious ratios of numbers and he postulated that since the seven known planets moved harmoniously the distances of these planets from the earth must be in the ratio of the lengths of the seven strings of the lyre. It is obvious that in the light of our present knowledge such speculation was futile. It was, however, man's first attempt at scientific thought, and it was tremendously stimulating. The difficulty with this method of thought lay first in the fact that it had inherent in it a *man-made faith that all things in nature were created for a purpose* as part of a well-ordered scheme. This naïve assumption came directly through man's attempt to read in nature actions similar to his own (i.e., it had an anthropomorphic basis). As far as man could see, his actions were governed by purpose, and in studying nature he assumed that it must also be arranged for a purpose. Had he resorted to experiments of a quantitative nature to test his theories and speculations, he would quickly have learned that there was no apparent purpose in natural phenomena, and he would have quickly changed his experiments in such a way as not to find an answer to a question "why," "for what purpose?" but rather to find an answer to a question "how" or "in what manner?" This experimental testing the Greek thinker did not do, for such a test involved the use of manual labor in conducting experiments. Now a freeborn Greek citizen might not be guilty of manual labor, for manual labor was the work of slaves. As a result the Greeks pursued the speculative method only. Thus it came to pass that, owing to the limited scope of man's sense perceptions, it was possible in 500 years' time to organize the universe accessible to man by these same five limited senses into a number of great philosophical systems so completely, that unless new experiences could be gained little further advance was possible. It is probable that this situation was responsible for the apparent stagnation of Greek learning at the time of the conquest of Greece by Rome.

It is clear from the nature and results of the above method that we cannot, as is frequently done, state that the beginnings of modern science, or even of electricity and magnetism, lay in Greece. It is true that about the first accurate records which we have of electrification of amber and jet by friction date from Thales of Miletus, 550 B.C., and that the first accurate mention of a magnetic attraction dates from that time. Doubtless these phenomena were well known long before this time and the Greeks added nothing but foolish speculation to this knowledge. It is often asserted by some that in reality nothing new exists in the world, and that the Greeks had already imagined all there

was to know of the universe. How futile and foolish such an assertion is will be seen in the later lectures. Not only did the Greeks make wild guesses, the majority of which are wrong, but even their philosophical systems are completely inadequate to cope with the analysis of the new universe of interstellar spaces and interatomic complexities now open to mankind through the extension of his senses. This extension of his senses, furthermore, came as a result of the scientific method of controlled quantitative investigation which arose in the sixteenth century and which was completely foreign to Greek thought. Occasionally, as in the atomic theory of Leucippus and Democritus, we have a speculation which was nearly like the accurately established picture of matter which exists to-day as the cumulative result of countless quantitative investigations. This *one* good "guess" is perhaps one of three or four of the hundreds of other theories that even remotely approach modern scientific conclusions.

In one direction, however, the Greek method was more successful. In the field of geometry, algebra, and astronomy, where physical manipulation was little required, the organizing power of the Greek genius was quite successful. Here the generalizations from observations of the ancients and further reasoning led to a real and valuable advance. Especially the development of geometry as summarized by Euclid in 300 B.C. forms a most valuable contribution to our knowledge which endures to this day, although the higher developments of even this branch in the non-Euclidean geometry remained for a new era to develop. In algebra, however, the progress was relatively slow because of the Greek system of numeration. Their numeral system was akin to that of the Romans, and the handicap which this gave to algebraic calculation can be seen at once if one tries to multiply the numbers LIX by XXIV, using the Roman notation. It remained for the Arabs at a later date to give to the world a rational system of numeration which made rapid and simple calculation possible.

Following the downfall of Greece, scientific advance nearly ceased for a long period beginning about 50 B.C. and ending only in the real birth of science in the Renaissance beginning about 1550. How this all came to pass is a long story. The Romans did not hinder Greek culture, and they even absorbed it. They were, however, too busy conquering, civilizing, empire building and governing to devote any time to the further development of Greek learning. With the downfall of Rome in about A.D. 410 to 450 the whole Western world was thrown into confusion. The Roman empire, with less than 5 per cent of its vast population forming the educated and governing classes, dissolved, completely submerging this small enlightened mi-

nority in the horde of uneducated barbarians and slaves that comprised the bulk of the empire. The whole of Western Europe was split up into a vast number of minor feudal states, all battling against each other. Thus, with little wealth, stability, or leisure, and with government in the hands of predatory feudal lords intent on war and plunder, even the remaining traces of Greek learning fell into oblivion. All man had time to do in this era was to struggle for existence in a warring world, with hunger and death on every hand. What spare moments were left were devoted to activities essential for the salvation of man's soul from eternal damnation. In the bitter conflicts between paganism and Christianity the zeal of the early Christians had in addition done much to destroy the Greek writings and manuscripts which still existed in the empire, so that Europe was by A.D. 600 deprived even of the opportunity of making available to itself the great advance made by the Greeks.

Fortunately for the world, the conquests of Alexander the Great and the Romans had carried to the Arab populations east of the Mediterranean the manuscripts and writings of the great Greek thinkers. These were promptly translated into Arabic, and thus preserved for Western Europe when in the eleventh century conditions had altered in such a fashion as to permit their further study. Through successful religious wars, the wealth and stability of the Arab population had reached such a stage that many Arabs could study and speculate on the work of the Greeks. While the work of the Greek philosophers was thus kept alive by the Arabs little was added by the latter to the Greek achievements. This was doubtless due to the nature of the Arabs, who were somewhat indolent, but probably more due to a different focus of interests. The science of astronomy to the Arabs was not as interesting as a development of the ancient Chaldean belief that man's destiny was written in the stars. Thus we see Greek astronomy diverted by the Arabs to astrology. Similarly, the Arabs actually dabbled with chemical reactions, again, however, with a utilitarian motive. This was either to transmute base metals to gold or silver, or else to find the philosopher's stone which would bring to its discoverer infinite knowledge and eternal youth. Alchemy became in the seventeenth century and later the father of chemistry, and we see in many of our present chemical terms such Arabian origins, as in the words *alkali*, *amalgam*, *borax*, etc. The most important contribution of the Arabs was, however, the Arabic system of numerals originally gleaned from the Hindus. These were developed further and led to an extensive application in the algebra whose development under the Greek system of numeration had been relatively slight.

The Arabic system consisted in the use of ten numeral symbols, 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. Numbers then were represented on a basis of multiples of 10 by the arrangement of the figures in the *order of precedence.* Thus, in the number 5342, the last symbol represents the units, the next the multiples of 10, the next the multiples of 100, or 10×10, the next the multiples of 1000, or $10 \times 10 \times 10$, and so on. Together with the use of the decimal point this gives a complete system of numbers whose multiplication, division, addition, and subtraction is exceedingly simple compared to the same process in the Greco-Roman system.

In the years following the downfall of Rome the Christian Church became a highly centralized and exceedingly powerful organization. With the revival of trade and the reduction of wars the wealth and power of the church still further increased. It then became necessary for it to develop schools and colleges where its future leaders could be trained. Furthermore, in the monastic organizations men could lead sheltered and peaceful lives of leisure devoted to study and worship. Between A.D. 700 and 1100 the large universities developed under the wing of the Church. The contact of the Western countries with the Arabs through the Crusades, through trade, and through the Moors in Spain, led to the introduction of the Greek and Arabic manuscripts into these institutions of learning. There thus developed great ecclesiastical scholars, men of high ability and great learning who quickly absorbed the early Greek culture and made it their own. Unfortunately, these men again scorned to work with their hands and adopted the philosophical views of the ancient Greeks rather than advancing these beyond the point at which the Greeks left them. They were still further hampered in any advance by the fact that the Greek writings had to be correlated with the writings of the Saints and the Scriptures. This led to a peculiar type of thinking which was called Scholasticism and completely dominated the Medieval universities. The influence of Scholasticism was such that the teachings of the Greek philosophers, especially Aristotle, completely dominated these institutions and prevented them from being the origin of the new science which arose elsewhere in the later days of Scholasticism. In fact the Scholastics fought the newer learning with a bitterness never since surpassed in the history of science. The extent to which these views dominated learning in the great universities is well illustrated when it is realized that even 200 years ago a student who doubted a single word of Aristotle could not graduate from Oxford, and the *science* of Aristotle was perhaps the most foolish of any of the Greek philosophers. In the time of Catherine de Medici, as Queen

of France, the learned medical men and scholars ***never touched their patients.*** The sick were treated by the ***barber surgeons***, indeed a very low caste of doctors as implied by the name. The Scholastics adhered to the Roman system of numeration long after the Arabic system had been adopted in business, in science, and by the world at large (i.e., from 1100–1200). The result is that even after Newton's day the scientific books were written in Latin, the language of the scholar, and to-day the chapter headings of the books are still in Roman numerals as are doctor's prescriptions.

There were, however, even in the early monastic universities occasional men who doubted Aristotle, and we find really able men, such as the Cardinal Nicholas of Cusa, Peter Peregrinus, and Roger Bacon experimenting with lenses, magnets, and electrified amber. The results of these investigations, however, are little available to us owing to their suppression inside the Church system, so that the true development of science was left to a new order of thinkers.

With the resumption of trade and the increasing wealth of the cities and states, and with the growing wealth and centralized power of ruling sovereigns and princelings due to the stabilization of government and trade, conditions for study and research in the arts and sciences were revived *outside of the ecclesiastical organization.* Many of the ruling powers had their court artists, litterateurs, and scientists who were free to pursue their studies whither their fancies led them without restriction. Furthermore, there was developing in the rich merchant classes a public which could support and appreciate art and learning in a manner at once dignified and satisfying. Cities and towns as well as the rulers founded colleges and universities in which there was an attempt to attract the leading thinkers and in which the latter could pursue their thinking and teaching without restriction.

It is therefore not surprising that beginning in about 1550 there began to arise groups of men who doubted the Aristotelian teachings and began to put these theories to experimental proof instead of arguing about them in the fashion of the early Greeks and the Scholastics. Among these perhaps the foremost was Galileo Galilei, whose researches really mark the beginning of our present scientific method of controlled quantitative investigation. Before Galileo there had however been a great thinker, who, despite the fact that he really used the older Greek method, had formulated an astronomical system whose influence on the future trend of science was prodigious because of the stimulus which it gave to experiment. This was Copernicus (1473 to 1543), a ***Prussian*** astronomer, who revived an ancient astronomical theory of

Aristarchus of Samos. Copernicus showed in a book that the existing astronomical data could be as well explained by assuming the earth and planets to revolve about the sun (heliocentric or modern theory) as by the theory of Aristotle that the sun and planets revolved about the earth on a series of crystal spheres along a group of curves known as epicycles. The book of Copernicus was published after his death, a fortunate circumstance, for it was of so revolutionary a character that he would probably have been martyred had he lived. The cause of the strife about this book lay in the fact that while Aristotle's theory made the earth the center of the universe so that the abode of the Pope (the Vicar of God on earth) was the center of the universe, the theory of Copernicus placed the Pope on an entirely insignificant planet revolving about the sun. This was rank heresy in the eyes of the Church.

Among the conclusions to which the Copernican theory led there was the one that the earth and planets were held in orbits about the sun by strong attractive forces. Such a view had a direct bearing on electricity and magnetism, for it led Otto von Guericke and Sir William Gilbert to investigate these phenomena, respectively, with the hope that the attractive forces of an electrified sulfur sphere for paper and light objects, and of a spherical magnet for iron, might account for the Copernican attractive forces. Both experiments failed of their initial purpose, but they led to the first *fundamental* advances in the fields of static electricity and magnetism of which we have a definite record.

The Copernican theory, however, itself remained no better than the Aristotelian theory until the telescopic investigations of Galileo in about 1600 showed the phases of Venus, the moons of Jupiter, as well as the rotation of the sun by means of sun spots. The real quantitative proof, however, did not come until nearly a century after the death of Copernicus, when Kepler, using Tycho Brahe's accurate observations, proved that the Aristotelian theory was totally incapable of representing planetary motion, while the theory of Copernicus sufficed to a surprising degree of accuracy if one assumed that the planets moved in ellipses which were nearly circular instead of in circles as Copernicus had assumed. The nature of the forces active was not determined until Newton published his " Principia " in 1687, and even then perhaps not successfully until the generalized theory of relativity in 1913 gave an even more general and more accurate description of gravitation.

However, as regards the true method of science as we know it, we can safely say that this method began to develop, as against the

Greek method, when Galileo put Aristotle's theory of the free fall of bodies under gravitation to test by dropping weights from the leaning tower of Pisa. From this time on the *method of controlled quantitative investigation* as the method for the study of Nature developed very rapidly in the hands of such masters as Galileo, Pascal, Torricelli, Huygens, von Guericke, and the Bernoullis, stimulated also by the more philosophical theories of Descartes as expounded in his " Système du Monde." In this book already was foreshadowed the philosophy of the new science.

Perhaps a word might be said concerning the oft-repeated slogan "controlled quantitative investigation." This method of study superseded the ancient Greek system of studying Nature which consisted of trying to determine "why" an event occurred. It endeavored to relegate to the background all preconceived notions as to "why and how nature *should* work" and proceeded to determine how Nature actually *did* operate in different cases by studying the phenomena quantitatively (so as to be able to reduce the relations observed to an algebraic equation). To do this in the more complicated phenomena required an experimental procedure in order to eliminate disturbing effects which obscured the result. This is the meaning of the word *control* in the expression used. As an example one might cite a study of the inclined plane. To make such a study we may take a cart of given weight which can be loaded with additional weights, place it on the inclined plane and fasten it by a cord to a scale pan of known weight on the vertical side of the inclined plane. Now an experimental study of the system above shows that the results are exceedingly irregular owing to friction. To *control* this troublesome factor the cord may be made to pass over a frictionless pulley at the top of the plane. The surface of the plane can be highly polished and the cart provided with smooth wheels. Further study would next show that the result is also influenced by the angle of the plane. One must therefore begin the study by *controlling* the angle of the plane, i.e., using a fixed angle for a first series of observations. With the angle fixed one may add weights to the cart and find the weights needed on the scale pan to balance the cart. These weights must next be recorded together with the lengths of the inclined plane and the vertical height of the pulley above the lower end of the plane. The weights can then be changed and a new balance found. A little further experimentation would then lead to the observation that the ratio of the total weights of the cart and the pan which produce a balance are in the same ratio as the two measured lengths. With this fact established one may then change the angle of the plane and

repeat the process. It will be found that the new angle gives a new ratio of lengths, and hence that the ratio of the weights changes in proportion to this new length ratio. A repetition of the experiment for several angles then permits one to write that the ratio of the weights of scale pan and of cart is equal to the ratio of the vertical height to the length of the inclined plane *within the limits of the experimental error.* Algebraically this is expressed by the equation:

$$\frac{W_p}{W_c} = \frac{L_v}{L_p},$$

where W_p = weight of pan and weights;
W_c = weight of cart and weights;
L_v = vertical height of plane;
L_p = length of the plane.

This at once establishes the law of the inclined plane as a result of a controlled quantitative investigation.

The success of this method was so great that already by 1687 the salient facts of elementary dynamics and statics were known. The formal recognition of this method as the future method of science was acclaimed when the newly founded Royal Society of London under Charter from Charles II of England established this method as the basis of its scientific mode of procedure in 1662. This date can in fact be looked on as one of the most significant dates in the history of science, for from that time on science, especially physical science, has progressed with an ever-accelerating rapidity.

Perhaps second in importance to the recognition of the future method of scientific study by the Royal Society was another feature whose significance can hardly be estimated. This was the beginning of the *publication* of its scientific proceedings. Up to this time the situation as regards publication and dissemination of knowledge had been as follows. With the Mohammedan occupation of the south-eastern portion of the Mediterranean the supply of papyrus (the Egyptian paper made of the bark of the papyrus plant pasted together) had been cut off in Europe and all the writings of any permanent sort were committed to parchment which came from split and specially prepared animal hides. Such parchment because of its oily texture cannot be used for printing. Thus all books and manuscripts were copied by hand, literally thousands of clerks and copyists being engaged in making such texts. Somewhere near 1456, to the best of our knowledge, in the workshop of one Fust in Mayence, Germany, a new process, printing, was invented for the manufacture of books. This

invention was due to simultaneous employment of three agents (paper, a proper ink, and block type) whose introduction was relatively novel. The use of paper originated in China in A.D. 105. The Chinese used a process for making a paper out of macerated rags suspended in water and run over proper flat strainers to dry. The process reached Bagdad in A.D. 795 and thence reached Europe about 1200. To the use of paper there was added by Fust the discovery of a suitable ink and the use of block type. The success of the method for reproduction of books was so great that in spite of some opposition there were by 1500 forty presses in action and eight million volumes had already been printed. Hence by the time the Royal Society was founded the development of printing had reached a high stage of development. The development of printing was in itself one of the greatest assets to the development of science, for it is quite certain that had it not been for the rapid reproduction by means of presses the then revolutionary works of Copernicus, Galileo and many others would have been suppressed by the Inquisition. Previous to the founding of the Royal Society, however, publication of scientific works had been a real problem. Most scientific workers were poor men and in order to have their works printed they had either to beg the money from royal and wealthy patrons by dedicating their books to them, or else they had to persuade some canny printer that the sale of their momentous work would bring the printer both money and fame. Hence the texts were long delayed in publication and poorly disseminated since modern publicity methods did not exist in those days.

The prompt publication of a scientific discovery as a short paper, within a few months of the completion of the work, in a magazine having wide circulation among scientific workers revolutionized the progress of science. Thus, through the general acceptance of the scientific method and prompt publication with efficient dissemination the rate of scientific advance was accelerated in an astounding fashion. To-day the scientific discoveries throughout the world are published in thousands of journals, and within three months after a discovery is announced by the pen of a worker it has spread throughout the scientific world so that the other workers may benefit by the newly discovered facts.

Despite the rapid development of new physical facts in the period around 1660 one factor was lacking for a really satisfactory progress. This lay in the fact that while the diverse phenomena of nature were being properly investigated and formulated there was relatively little basis for coordinating quantitatively and theoretically the results in various fields and hence in relating phenomena in the fashion in which

they should be related. That is, it will be noticed that many phenomena, to-day closely correlated, were treated as separate phenomena for lack of a consistent system of units and a uniform mode of relation and expression. This will be seen in the work on electricity and magnetism preceding 1750.

At the age of twenty-one a young English physicist, born in the year in which Galileo died, had the keen insight to arrive at a general formulation of the laws of mechanical action. This man was Isaac Newton. Owing to certain discrepancies between theory and experiment, which were used as the basis of a crucial check of his method of formulation, Newton did not publish his general scheme until 1687. The discrepancy as it turned out lay in an error in the measured radius of the earth and not in the theory, so that when a new survey of the earth was completed it was found that the theory and observation agreed as well as could be expected and Newton's "Principia" appeared in 1687. The theory was readily acclaimed in England but it took some years before the world as a whole accepted it and the methods of generalization to which it led were widely applied. The general acceptance can be seen by the introduction of the quantitative method in electricity and magnetism by Cavendish and by Charles A. Coulomb in 1780–1789 during what we shall call the static period. These investigations give an excellent example of the importance of Newton's contribution, for by means of the forces as defined in the Newtonian system Coulomb was enabled to formulate a definition of unit electrical quantity and unit magnetic pole, which could be so related with other mechanical phenomena that the whole of electricity could be expressed in simple and measurable mechanical terms.

The basis of Newton's generalization lay in the definition and formulation in a precise and measurable manner of two quantities—inertia, or mass, and force—by means of Newton's three famous laws of motion. The remarkable power of these laws and their influence on our thinking is little realized even by the student of physics. However, a survey of the development shows that from 1700 on until 1900 the whole development of physics depended on the interrelating of various phenomena with each other by means of the Newtonian mechanical system. This period of development of the Newtonian system can be well divided into about three sub-periods. In the first, while correlation and test of the Newtonian laws as applied chiefly to mechanics was going on, other phenomena such as electricity, magnetism, heat and light were studied only as isolated and widely separate subjects in a qualitative way. From about 1750 on we see a general development involving the quantitative development of

these fields, the new forces and experiences being measured and defined in terms of mechanical actions as indicated by Newton's laws. This led to a wealth of quantitative material so related that it was available for far greater generalization than could have been otherwise dreamed of. From 1850 on we see these remarkable generalizations taking form and reaching their height during the last years of the nineteenth century. These generalizations perhaps began with the atomic theory and the doctrine of conservation of matter (Lavoisier, 1780–1790) but more so as a result of the discovery by experiment and generalization of the principle of conservation of energy. It is of interest to note that Joule in his experimental test of this principle used the mechanical calculation of electrical energy made possible by the Newtonian definition of electrical forces. The principle of conservation of energy applicable to all systems enunciated by Mayer and Helmholtz was followed in 1850–1851 by the second law of thermodynamics enunciated by Clausius and Kelvin, and in 1865 by the Maxwellian formulation of electrodynamics. These generalizations led to a remarkably complete description of the whole of the physical phenomena then known in terms of a beautifully coherent system based entirely on the fundamental axioms of Newton. It was then not surprising that leading physicists in 1890 should have said that our knowledge of physics was complete and satisfying and that further discovery lay only in the fourth significant figure. This complacency of the physicist was soon to be completely shattered, for already in 1895 x-rays were discovered, and with them came untold grief for the complete and satisfying Newtonian universe.

The effect of Newton's laws was even more far-reaching than the mere mathematical formulation of all nature into an apparently consistent system. These laws and their implication colored the whole philosophical attitude of the physicist. We have seen the manner in which the Greek question "why" relative to Nature was modified to a question "how," limited to a quantitative description of a phenomenon. Already in the time of Galileo, Descartes developed a completely new philosophy based on this newer aspect of investigation. This was the well-known mechanistic picture of the universe which played so important a rôle among the great French liberal thinkers who set the stage for the French Revolution. Had the answer to the question "how" as placed by Galileo and the earlier workers remained as it was we would perhaps have been spared the great difficulty which faced physics in the period from 1900 to 1927. Unfortunately, the question "how" asked of Nature did not remain just a question

answered by an equation as derived by a specified controlled quantitative investigation. Instead, with the adoption of Newton's laws it became a definitely modified question "how." For following the appearance of the "Principia" the question was always "how is this phenomenon explained or described *in terms of Newton's laws?*" As soon as a new phenomenon could be reduced to equations of a Newtonian character and thus related to all the rest of physical knowledge the physicist went away with the happy feeling that the phenomenon was "explained." The term "explanation" used, however, always implied more than the physicist should have admitted, for he should have added to the word "explained" the limitation "in terms of Newtonian mechanics." This subtle error led to very naïve blunders in physics, for towards the end of the nineteenth century many of the most distinguished physicists attempted to "explain" the action of light in the ether and in substances by means of rough mechanical models of weights, springs, and frictional elements, whose mechanical equations roughly approximated the electromagnetic wave equations but whose action was fundamentally utterly different. Raised in the atmosphere of the Newtonian mechanical concepts the whole of physical thinking was colored by these concepts, and implied in these was an *absoluteness* and a rigidity of thought and formulation which boded ill for the future.

Already by 1900 discrepancies between the energy radiated by hot incandescent bodies and the theoretical deductions of the intensity of radiation on the basis of classical mechanics and electrodynamics were discovered, and the birth of the quantum theory of light emission in the hands of Max Planck caused some concern. Then in 1905 came Einstein's special theory of relativity, which threw down the gauntlet to Newtonian mechanics as initially formulated. This was followed by Einstein's general theory of relativity, which received four striking experimental verifications by 1919. In this development it was seen that Newton's laws were but a first-order approximation to far more accurate and general laws, and that Newton's laws broke down completely when bodies began to move with velocities comparable with that of light. With this theory the crude mechanical ether based on Newtonian explanation and the theories of Maxwell vanished completely. All that remained were accurate mathematical formulations of a very general nature which more accurately represented Nature than had Newton's concepts.

Next came indubitable proof of the quantum nature of light in the photoelectric emission (i.e., emission of electrons by light and x-rays) followed by the complete breakdown of classical electrody-

namics in the exchange of energy between rapidly moving electrical charges and atoms. These results came as a consequence of the beautiful experiments of Millikan, and Franck and Hertz. They were most disturbing, but the crowning difficulty came with Compton's discovery that short wave-length light waves (x-rays) collided with the free minute spherical electrons, as conceived on the Newtonian viewpoint, as if the light waves themselves were bullets. This was a most serious difficulty, for all our knowledge of light indicated it to be an electromagnetic wave motion, and x-rays were certainly, like light waves, such electromagnetic waves. Something was obviously radically wrong somewhere.

In 1926 to 1927 a group of several mathematical physicists independently began to solve the difficulty. These men were Heisenberg, de Broglie and Pauli, Schroedinger, and later, Dirac. They showed that our naïvely pictured spherical Newtonian electron should be replaced by a complex set of mathematical equations which depended on the state of the electron. These equations enabled the electron, under circumstances where it reacted like a particle, to be represented by such an analogy, and in circumstances where it encountered and reacted with light waves to act like a complex wave motion. With the experimental proof of the correctness of these view by Davisson and Germer, Thomson, Rupp, and most elegantly and completely by O. Stern and his collaborators the faith in the new revolutionary developments was established. With this confirmation of the wave mechanics, as our new method of physical formulation is called, the physicist has been cured of his naïve error in the blind acceptance of a mechanical universe based on Newton's laws. Exactly what form the new philosophy of physics will actually take one cannot predict. All that can be said is that "explanation" in physics will no longer be limited by the Newtonian law, and that in its place the phenomena will be described as closely as possible by the accurate and involved expressions of the wave mechanics, relativity, or similar formulations, with the probable added restriction that even these may be only approximations to still more complete and accurate expressions yet to be discovered. What we can then hope to do is to measure and describe as best we can by means of the method of controlled quantitative investigation the phenomena of Nature and to formulate these in the most convenient general systems of representation which we have at hand. These will enable us to predict with some probability of success the behavior of our system under various conditions, and will enable us to further investigate and control it. They will also enable us to utilize the phenomena in one way or another to improve the condition of man on

earth by furnishing new comforts, luxuries and conveniences as to material civilization.

Before turning to the detailed development of electricity and magnetism which can be outlined against this general background a few general comments might be made. It must be seen from what has gone before and what will follow that in each period the possible discoveries much depended on the general state of knowledge at that period. Thus it was to be expected that the discoveries of Coulomb and of Cavendish should have come within the period say 1730–1800, and that Maxwell's brilliant generalizations must have been due to arrive after the careful quantitative discoveries of Faraday and at about the same period as the generalizations of Helmholtz, Clausius and Kelvin about energy. It is thus seen that a discovery is made at a time when conditions are ripe for the discovery, or in other words that a given discovery is "in the air" at the time when it is made. It is thus merely a chance that the discovery is made by a given man, and many a discovery is made independently and nearly simultaneously by several men. Cases are on record (such as the discovery of absolute 0 of temperature by Amonton in 1700 or of the mechanical nature of heat by Davy, Rumford and others in 1800, perhaps a hundred or in the second case thirty years before the accepted date of discovery) where a discovery is made by chance before the time is ripe. In general, however, the man making an important advance in a given field is the man who is most actively engaged in an intelligent research in the field at the time when the stage is set for the discovery. An excellent example is the discovery of x-rays by Roentgen, of electromagnetic induction by Faraday, or of the electromganetic waves by Hertz. Sometimes the great discoveries are made by accident by relatively inferior men, as perhaps was the case with the discovery of the magnetic effect of a current or of the phenomenon of radioactivity. Thus the greatness of a scientific discovery does not necessarily indicate genius in the discoverer. The test of greatness rather lies in a man's seeing a relation clearly long before it is seen by others and working it out accurately and completely before it is handled by others. Perhaps the truest test of ability is the successive discovery of several important and far-reaching facts by a single man. On this criterion men like Galileo, Huygens, Newton, Ampère, Thomas Young, Faraday, Clausius, Maxwell, Rutherford, Bohr and Einstein could most easily be classed as the greatest leaders in physical thought. It might be added that to-day with the remarkable interest, facilities and number of workers in scientific fields, together with the rapid publication and dissemination of knowledge the competition is so keen and the

rate of advance is so rapid that it is almost impossible to credit great achievement to any one man. There is hardly an important problem attacked in which at least two men are not working along closely parallel lines and in which the results are not published nearly simultaneously. The recent theoretical and experimental proofs of wave mechanics furnish an excellent example of this, for it is almost impossible to find any unanimity of opinion among scientists as to who deserves the chief credit for these spectacular advances.

The appalling rate of increase of physical knowledge might lead one to ask whether this accelerative increase of scientific production will go on indefinitely. The acceleration is due to the fact that the more facts there are known, the more analogies we have to build on, and the more facts to correlate, the more rapid advance will be. It is, however, conceivable that the time is approaching when the new facts found will accumulate so rapidly that the single human mind will not be able to hold enough facts to make use of the new correlations. The result will be that the rate of gain of knowledge will reach a steady state, the limitation being that produced by the incapability of the human mind of grasping more than a certain number of facts at once.

2. THE DEVELOPMENT OF ELECTRICITY AND MAGNETISM

With the above remarks we may proceed to a more detailed study of the development of electricity and magnetism. To facilitate the study we can consider the development as divided into seven general periods. These begin approximately in the great period of development of the scientific method starting with Galileo. Inasmuch as electrical and magnetic phenomena are more complicated to study and to observe than the simpler mechanical problems attacked by Galileo it is not strange that these fields should have begun somewhat later. In fact it is quite easy, neglecting the development of electricity and magnetism in the Greek period which was trivial, to start our history with the first extensive qualitative investigation of magnetism. The seven periods into which electrical and magnetic advance may be divided are as follows:

a. Static Period, 1600–1799, and divided into:

   1. Qualitative Period.
   2. Quantitative Period.

b. Electrical Current Period, 1799–1831.

c. Electrotechnical Period, 1831–1865.

d. Systematic Period, 1865–1897.

e. Atomic Period, 1897–1919.
f. Quantum Theory and Relativity Period, 1919–1927.
g. Wave Mechanics Period, 1927—.

a. The Static Period, 1600–1799.—This period begins in 1600 with the publication of the first experiments on magnetism, generally attributed to Sir William Gilbert, physician to Queen Elizabeth, in his book " De Magnete," and extends through the early qualitative experiments on electricity through the quantitative measurements of frictional electricty under Coulomb up to the publication of Volta's discovery of the Voltaic pile in 1799. The discovery of the Voltaic pile marks the beginning of a new period in electricity, i.e., the period characterized by the development of the study of electrical currents by their magnetic effects as a separate type of electrical phenomena. The two types of electricity while actually recognized as of the same origin were however not clearly related until the relations of the electrostatic and electromagnetic phenomena were established by the work of Weber and Maxwell in about 1845 to 1865.

While in general Sir William Gilbert's book is acknowledged as one of the forerunners of the new method of scientific investigation and contains the essential qualitative phenomena of magnetism, excepting those dependent on electrical current, the recent evidence seems to indicate that Peter Peregrinus in 1269 in his book " De Magnete " had anticipated several of the important discoveries of Gilbert. Peregrinus was one of the Scholastics and the early development of the true scientific method of research in this period must be classed with the early experimental work of Archimedes, 287–212 B.C., as quite anomalous. Peregrinus made a large spherical magnetic lodestone and by means of small magnets showed that this sphere had two poles and that lines of force like meridians radiated from the poles. He showed that iron had magnetism induced in it by the proximity of a magnet, and that like poles repel and unlike attract. He also explained the earth's magnetism in terms of his sphere. He further showed that broken magnets are still magnets, and stated his belief that magnetism of bodies on the earth is due to the earth's field.

Gilbert's book appeared in 1600 and antedated Galileo's book on the methods of the new science. It was quoted by Galileo and Kepler, and Hoppe regards Gilbert and not these latter as the founder of the new scientific method. It is also claimed that Descartes was very much influenced by this book. The investigations leading to the book were the result of Gilbert's desire to prove the Copernican theory. Gilbert extended the discoveries of Peregrinus and showed the dip

of the magnetic lines. He found that heat destroys magnetism and that jarring a steel body in a magnetic field induces magnetism. He showed that magnetization seems to be increased by extending the poles so as to approach each other. He distinguishes the difference between the attractive forces on magnets in an inhomogeneous field and the directive action on the magnet of a homogeneous field. He characterized declination, and dip or inclination. He attributed the earth's magnetism to iron ores in the earth, and showed again that magnetization does not increase the weight of a magnet. The development of magnetism from the time of Gilbert until the measurements and definition of magnetic pole strength as a result of Coulomb's experiments in 1785 was practically at a standstill. This is not strange, as Gilbert had probably developed magnetic studies qualitatively as far as was possible until the Newtonian mechanics made further quantitative study possible. In the second part of his text Gilbert treated of electricity. The Greeks had only inaccurately observed electrification of amber and some other unknown substance, lynkurion (probably jet). Even the electrification of the second substance had been forgotten. In this part of his book Gilbert calls the phenomenon electrical, and distinguishes between substances which are electrical and which are not electrical, i.e., which cannot be electrified by the hardest rubbing. He uses his idea of the magnetic dip needle to measure and test for electrification, thus inventing the first crude electroscope. The effect of air moisture on electrical phenomena was noted as well as the discharge of a charged body by a flame. He also points out clearly the difference between electricity and magnetism.

From this point on until 1671 there was little or no experimental progress in electricity, and the theoretical speculation of the workers Cabeus and Kircher was ludicrous. The next experimenter in electricity was, like Gilbert, led to the investigation by the Copernican theory. This was Otto von Guericke, the Mayor of Magdeburg, who is best known for his work on air pressure. Von Guericke's progress was due to his development of a new electrical machine. He cast a sphere (i.e., like the earth), of sulfur, mounted it on a lathe and rubbed it with the dry hand. The electrification so produced was large enough to give much greater latitude to experimentation. · He showed the attraction for light objects including water drops which shattered before they struck the sphere. He showed that the attraction was transmitted over six feet of linen cord. If the light objects attracted to the sphere were allowed to touch it they were repelled until they had touched another body. A feather brought near the sphere showed

electrification. If the feather was touched and the hand removed it was repelled and the fibers stood straight out. On touching a neutral body the feather became limp again until it approached the sphere without touching. This was the first observation of induction. He also observed the discharge by a flame. He next observed the noise, and the light of the sparks at night. Most of these fundamental discoveries were forgotten and attributed to others. Conduction was later rediscovered by Gray (1729) and in the same way induction was attributed to Du Fay (1733), while the discovery of the sparks was attributed to Wall in 1708. This perhaps is unjust, but it must be remembered that at these later dates the rediscovery of these facts resulted in a rather exhaustive study and, in general, a correlation with other effects which served to fix them in mind. However, even in interpretation, Guericke was well ahead of his time. He noted that the attractive and repulsive forces varied with the distance and fell to 0 at greater distances, and he actually differentiated the concept of force from that of material bodies. In fact, in this direction, he was far ahead of the later workers, even including Newton, whose electrical speculations were rather weak.

Following Guericke, further advances were made by Robert Boyle, who was aware of Guericke's work. Thus, while Boyle confirmed some of Guericke's experiments he found little new except that the attractions occurred in the absence of air as well as in air. To Boyle the use of his new air pump was always a source of great amusement. It is not strange then to find Boyle's name coupled with early investigations in many fields in which the phenomena were tried *in vacuo*. Boyle placed everything from bells and rubbed amber to mice under vacuum and studied the effects. At times the results were of considerable importance and other times more messy than interesting.

By 1708 Wall had observed long sparks drawn by his finger from amber rubbed with wool and had definitely compared them to thunder and lightning. This early guess, as to the identity of lightning and the electrical spark, received unsuspected support at about this time by the discovery in changes of magnetization of steel bodies subjected to the proximity of lightning discharges. At the same period we find Hawksbee's study of the first "electric light" as it was called. In 1705 Hawkesbee found that if the sealed end of a barometer tube filled with mercury was tilted back and forth so that the mercury moved inside, a peculiar faint light moved up and down in the tube with the mercury. He showed by a brilliant series of experiments that it came from the frictional electrification, and proved that electrical discharges outside produced the glow observed inside.

In 1729–1731 Stephan Gray repeated Guericke's experiments and found that under proper conditions the electrification from the sulfur sphere could be conducted some fifty feet or more by a linen thread and detected. His success lay in the use of insulating silk threads in suspending his long moist linen string, an idea suggested by Wehler. The idea of insulation being brought to his attention he insulated a boy on a proper device and found that he became electrified by the approach of a charged glass sphere (Gray used glass instead of sulfur), thus rediscovering induction. From then on boys became splendid subjects for experimentation of this sort. The discovery of insulators led to the use of horn and other resinous substances for insulation. The real contribution of Gray was thus less the rediscovery of conduction, than the discovery that electrification could best be studied by the use of insulators. He was therefore led to distinguish materials as conductors or non-conductors (insulators) of electricity. The non-conductors were found to be electrified by friction. Gray also distinguished electrical brush discharge from sparks, and that the surface rather than the mass of a body was of importance in electrical phenomena.

Contemporary with Gray's work in England there was perhaps a more distinguished and certainly more brilliant investigator in France engaged in work along similar lines. This was Du Fay. He repeated and extended Gray's experiments, suggesting that probably even metals could be electrified by friction if they did not lose their charge by conduction. He believed conductivity to be possible over indefinite lengths. For experiments he also used boys and improved somewhat the insulation suggested by Gray. He used the repulsion of two light metal foils as a test for electrification, a more sensitive modification of Gilbert's electroscope. With a further improvement of the electroscope he discovered that resinous bodies rubbed with wool repel each other, but are attracted by vitreous (glassy) substances rubbed by silk, the glassy substances repelling each other also. This led at once to the recognition of the two types of electricity, vitreous and resinous, and that like electrification repels while unlike attracts. This discovery, together with induction, and the discovery of the Leyden jar in about 1745 led Du Fay to postulate that there were two weightless electrical fluids, resinous and vitreous, and that neutral bodies carried equal amounts of both. This theory was assailed by Franklin, who opposed the one fluid theory which was finally adopted.

In the meanwhile the crude electrical machines of Guericke and of Gray were improved first, by replacing the hand with a leather glove, next by using metal collectors, and then by using Leyden jars

for accumulating the charge. By 1755 the use of rotating glass discs was already introduced so that by 1786 quite satisfactory sources of static electricity were available.

In 1745 the prelate von Kleist in Pomerania discovered that the electricity from a charged sphere could be stored in a bottle if it contained copper wire or mercury and the latter were brought in contact with the machine while the bottle was held in the hand. From Berlin, where this was shown before the Academy, the Burgomaster Gralatt of Danzig took his cue and attempted to repeat the experiments. There were some difficulties at first and the news did not get beyond the German border. Meanwhile in Holland von Musschenbrock had a year later, 1746, independently made the discovery so that it reached France through the Abbé Nollet before von Kleist's earlier discovery was heard of. Naturally, the peculiar power of storing electricity in a bottle or a jar coated inside and outside with a metal coating, which was not in contact, became associated with von Musschenbroek and his native town of Leyden, Holland, and accordingly was perhaps erroneously called the Leyden jar. At about the same time Gralatt found the weakness in Kleist's experiments, and substituting metal or an outside covering for the hand he got results similar to the results in Leyden. The Leyden jar, however, does not seem to have led to any great advance comparable with the interest which it aroused. Its chief value was its ability to accumulate larger quantities of electricity than was previously possible and that it brought home clearly once and for all the idea of the electrical circuit. It further played an important rôle in the controversy between Du Fay and Franklin about the nature of electricty. Franklin believed that neutral bodies carried a normal amount of electricity. Any excess of vitreous (positive) electricity caused a body to be positively charged, while a lack of the weightless positive electrical fluid left bodies with a lack of electricity (i.e., a negative or resinous charge). No decision between the one and two fluid theories could be reached at that time but public opinion gradually swung over to Franklin's point of view after the time of Faraday. To-day we know that both views were only partly right.

In 1746 Le Monnier measured the speed of travel of electricity in a wire and concluded that it was very high. During this same period Franklin was beginning some of his electrical experiments. At first his experiments were conducted as parlor tricks to amuse his friends. This had been the fashion in the Paris salons of the period. Later, however, he carried on his experiments more in earnest. His studies occupied themselves largely with the nature of lightning. It was sus-

pected that raindrops carried an electrical charge, and based on the assumed correctness of some inaccurate observations, Franklin proposed a theory of thunder showers. He was convinced that lightning was an electrical spark on a large scale. He further made use of an idea by then well known that sharp points discharged electricity and proposed to neutralize the electricity in clouds which caused lightning to strike by the use of sharp points on grounded conductors. This is one of the superior methods to-day in protection against lightning. He proposed to set a man on an insulated stool and with a sharp, pointed conductor sticking out through the roof of a hut detect the electricity discharged by the rod to the cloud by observing the resulting residual charge left on the man by the sparks from the point. The proposal was made in 1750, and in 1752 d'Alibard in Marly, France, tried the experiment. The same experiment resulted in the death of Richmann in Petersburg, who was electrocuted by too heavy a discharge. In the meanwhile Franklin tried the still more dangerous feat of drawing electricity from the clouds by a kite string. As long as the string was dry he got no charge but as the rain moistened the string he drew down enough electricity to give sparks and to charge a Leyden jar which then could be discharged at will. He further proved the existence of the electrification by utilizing the charge on the jar to ring the first electrostatic electrical bell described in Figure P 21, which had been discovered by Gordon in 1741. While Franklin's proof of the identity of lightning and electricity produced by friction was not of great theoretical importance for the study of electricity, it was perhaps, in its psychological effect, one of man's greatest triumphs. For it showed that by the new scientific method one could study and in a measure control one of nature's most splendid and at the same time, terrifying phenomena. Thus the much-feared thunderbolts of Jupiter were reduced to a mere laboratory phenomenon, a fact which did wonders in liberating mankind from the superstitious dread of Nature, and showed him that by perseverance he might himself aspire to a real conquest of Nature. The proof given by Franklin further led to the development of the lightning rod, which in its essence has not been improved to-day, and which has saved many a home from destruction.

Around 1750 the development and general acceptance of Newton's mechanical system had begun to bear fruit and we find attempts made to measure electrical effects *quantitatively* by means of the mechanical action of electrical forces. The work of Waitz, 1745, of Ellicott and Gralatt, 1747, and of Abbé Nollet was directed to this end in attempting to balance the repulsive forces of charged bodies

by gravitational forces. Why it was that these experiments did not lead to a fruitful result is not clear. Perhaps the mechanical analysis was beyond the ability of the workers and perhaps the control of the experiments was not sufficiently developed. However, in the periods following these attempts there were developed several devices by which the repulsive forces could be measured. One of them, due to Henley, was the first quadrant electrometer, an instrument now used to measure small potential differences. The equations involved were, however, too complicated to lead to any fundamental laws. Volta, in 1781, enclosed the light straw leaves of the early electroscope in a vessel free from air currents and used a scale to measure their deflection. He therefore perfected the electroscope. Here again the theory was too complex to lead to any progress. In 1787 Bennett used gold leaves instead of straw for this purpose, thus perfecting what is virtually the modern form of the gold leaf electroscope. The electroscope was also improved by Bennett to be used with the potential multiplier described on page 210. The rôle of this instrument was important early in the study of the galvanic cell as it enabled Volta to prove the identity of his electrification from the Voltaic pile and static electricity.

The really successful first quantitative study of electricity published was due to Charles A. Coulomb. That he should have succeeded where others failed is not strange. He had been led in 1779 through a study of the deflection of a compass needle to a study of what is known as the torsion balance. This instrument consisted of a rigid bar suspended horizontally at its middle by a vertical cord or wire. He studied the twist produced in this suspension by means of a given force applied to the end of the lever arm (i.e., the deflection or twist produced by a known torque, see section 14). He measured the deflection or twist by the displacement of a beam of light reflected from a mirror placed on the vertical axis of the bar. Having accurately derived the mechanical laws of his torsion balance (i.e., as it depended on the length, diameter and material of the suspension) he had a new mechanical measuring device whose mechanical properties in terms of forces and distances he could discuss and analyze clearly. In 1785 he turned this device to the measurement of electrical and magnetic forces. In the electrical case he hung a charged body at one end of the horizontal rod and tested the twist produced by an equal charge on another body of the same size placed at various distances from the charged end of the rod. The deflections read off could at once be interpreted in terms of forces. He quickly discovered, as the result of a series of experiments, that the force between electrified

bodies could be expressed in terms of an equation $f = k\frac{qq'}{r^2}$, where k was a constant dependent on the units of force, r the distance between the centers of the two charged bodies, f the measured force and q and q' magnitudes expressing or characterizing the *degree* or *state* of electrification of the spheres. This at once defines electrical quantity, for if q and q' be made equal by touching the two bodies then $q^2 = \frac{fr^2}{k}$, and $q = \sqrt{\frac{fr^2}{k}}$. The unit of charge or electrical quantity is then at once obtained from conditions that make $f = 1$, $k = 1$, and $r = 1$, for then q will be 1. This work of Coulomb's was the first published investigation which gave a definition of unit electrical quantity (a similar law was found by Coulomb for magnets where q and q' were replaced by m and m' representing the magnetic pole strengths of the magnets). It is said that Priestley, in 1767, had guessed that the force between charges should vary as $\frac{1}{r^2}$, and Michell in 1750 following the experimental results of Musschenbroek had deduced the law for magnetic forces. Independently, Tobias Mayer, 1769, and Lambert in 1766, had claimed to have discovered the same law. However, the complete and accurate experiments of Coulomb bring him the credit for the work which he apparently justly deserves, for it was his experimental work that definitely established the law. Years later Maxwell in 1879 published the notebooks of a queer scientific recluse named Cavendish. Cavendish, a man of wealth and leisure, carried on experiments on electricity and magnetism and on many allied subjects. For some strange reason he had seen fit to publish but two of his works, one of them being a direct determination of the gravitational constant as a result of the attraction of large masses for smaller ones suspended in a torsion balance in 1798. In these notebooks it was found that perhaps six to ten years before Coulomb Cavendish had discovered the law of electric attraction (i.e., that f was proportional to $\frac{1}{r^2}$) by means of a very clever piece of reasoning based on experiment. Since these results were buried and made unavailable to science the credit really goes to Coulomb and we can set 1785 to 1789 as the proper date for the quantitative beginning of electricity and magnetism.

With the discovery of the law of force the whole field of static electricity could quickly be coordinated with the already highly devel-

oped body of information based on Newtonian mechanics. Thus in a relatively short period the mathematical physicists Biot, La Place and Poisson developed the field of static electricity quite extensively. These men were active in this field towards the end of the eighteenth century and until the end of the first third of the nineteenth century. It was, however, only in the systematic period, together with other great laws of physics that the finer definitions and advances were made.

b. Current Period, 1799–1831.—The second period of development begins with the discovery of the Voltaic pile or galvanic cell as a new source of electricity in about 1799 and extends to the time of Faraday's discovery of a third source of electrification in 1831. The period while containing the further development of static electricity was entirely colored by the new source of electricity and the further knowledge to which it led.

The train of circumstances which led to the discovery of the Voltaic pile is most interesting. Even in the time of the ancients the power of certain fish to inflict a painful shock of some sort when touched in the appropriate places (top of the head and underside of the body) had been well known. One of these was the electric or torpedo ray and later the electrical eel. It appeared that Adamson brought one of these electric fish, *Silurus electricus*, to London from Africa in about 1751 at a time when the discharges from the Leyden jars were being frequently demonstrated in the salons and academies. The analogy of the sensations produced by the fish and the jar at once led to the belief that both effects were of the same origin and the fish were called electrical. In 1773 Walsh showed that the shock was obtained only by contact at certain points. He then isolated the electrical organs in the head of the ray and obtained further proof that the effect was in close analogy to electricity in its behavior. It was not however until 1837 that Linari showed the electrical nature of the effect by charging an electroscope from the fish. However, this work excited a great deal of interest at the time. Among the men interested in this animal electricity was an Italian physiologist, Galvani. It happened in 1780 that while he was demonstrating a nerve-muscle preparation of a frog's leg another group of students was working with a static machine nearby. One of Galvani's students noticed when the nerve was touched by a metal instrument the muscle twitched every time there was a spark in the static machine. It was quickly found that the electrical discharge *directly* caused the effect and the fact was still further corroborated when the nerves were hung out on wires exposed to the effect of electrical storms or clouds, as Franklin's work led one to expect. It was further observed that

neither the electrical machine nor the clouds were needed. A twitch in the leg could always be produced by touching two dissimilar wires together at one end and placing the other ends in contact with the nerve. Had this been the sole effect Galvani might have been led to reap the reward of his own discovery. However, it was also found that a piece of iron or copper alone would produce the twitch under some undefined conditions, if the two ends touched the nerve. This fact disconcerted Galvani and he believed the effect was due to a discharge of the electricity from the nerve muscle system through the wire, or some such electrophysical action. One or two of the many subsequent observers still believed that the metals produced the effect, but most workers, including Volta, a friend of Galvani, inclined to Galvani's viewpoint as late as 1791. Volta, however, continued his investigations, using a different set of tests. He repeated some earlier studies of Sulzer dating from 1754, in which he tested the effect of these wires by applying them to the tongue, or to the lips on the one hand and a moistened contact above the eyes on the other. The first produced a taste sensation, acid-like at one metal, alkaline at the other. The second stimulated light sensations in the eye. With these tests Volta quickly classified the different metals into groups as regards their effect. He also found that a wire of a single substance was ineffective *unless the two ends were rendered different* by different treatment (i.e., one end hot, the other cold; or one end hardened, the other soft, etc.). By 1794 he had definitely placed the seat of the action in the metals. He then used the multiplier and electroscope and showed that contact of the two metals actually produced electrification (1796). Between 1796 and 1799 Volta succeeded in showing that the effects occurred when moist rags with conducting solutions of salts were placed between the two metals, and that the effects became even more powerful if a series of alternate plates of dissimilar metals (e.g., Zn and Cu) were placed in a pile with wet cloths between them. This was the famous Voltaic pile or galvanic cell. Volta sent a report of this discovery in 1800 to Banks in England who turned it over to Carlisle for publication in the Royal Society. Carlisle kept the note and together with Nicholson repeated the experiments and published them as their own experiments before they published the letter of Volta. This cheap trick availed them little, for Volta's discovery had reached the world through other channels and Nicholson and Carlisle have throughout history been branded as scientific thieves and cheats, an unenviable reputation indeed for a little false glory. At once the whole scientific world was constructing Voltaic batteries and studying the new effects possible by their means.

That the discovery was "in the air" even before the publication is evidenced by the fact that a German chemist, Ritter, had observed in 1799 that the two wires of different metals touched together at one end caused an evolution of gases when the free ends were introduced into a drop of water. Even in 1796 Ash had noticed the electrolytic corrosion of metals in contact in a salt solution. Ritter, by 1800, showed the gases to be oxygen and hydrogen and thus first observed the phenomenon of electrolysis. He also observed the deposition of copper from solution by the two wires. He connected this with the physiological action and developed a type of galvanic chain or cell. By 1801 Ritter had developed the idea of a series of solution tensions (page 150). This potential series of elements as regards their activity in cells was later also independently discovered by Volta and made the subject of a long series of experiments. It is known as Volta's electromotive series. Thus the peculiar electrical activity of nerves had done their work and the physicist had a new group of phenomena with which to play, and to discover even more remarkable facts.

Already, by 1808, Humphry Davy showed that beyond the simple electrolysis of salts in solution, discovered by Ritter and others, the supposedly elementary alkalies of Lavoisier could in the molten state be electrolyzed. Thus he discovered sodium and potassium. In 1810 the actual transportation of the ions in solution in animal tissues had been looked for by Wollaston and in 1816 Porret's experimental observations of this transport were first published. Other researches led to an improvement of the Voltaic piles or galvanic cells and a clearer understanding of the shortcomings of these batteries, including such effects as polarization, etc.

In another direction, however, investigations were even more fruitful. Since the early observations of lightning discharges on magnets various attempts at finding a relation between electricity and magnetism had been made, but without success. In 1812 Oersted had stated that he believed that "the electrical forces under circumstances where they occur tightly bound should be able to cause an action on the magnet as a magnet." Just what was meant by this "tightly bound" condition is not clear, and it indicates a very unclear notion of the phenomena which he was looking for. Oersted believed, however, that the conditions in an incandescent platinum wire short-circuiting a galvanic cell represented the supposedly tightly bound state. In 1819 he placed such a wire over a compass needle and observed an effect on the needle, albeit weak and uncertain. He repeated the experiment with a cell of 20 elements and got a marked deflection for a wire $\frac{3}{4}$ inch above or below the needle, the deflection being 45°. The results of his

studies Oersted published in a Latin monograph sent to all outstanding scientists on July 21, 1820. Oersted noted therein that the deflection was dependent on what we *now* term the current strength, and that a movable current circuit is deflected by a fixed magnet pole. Oersted found that the effect was more increased by the increase in area of the battery electrodes than by the number of batteries in series (i.e., that with the poor batteries of those days small internal resistance was more important than increase of potential). Gay-Lussac and Arago shortly thereafter showed that those "currents" could also produce magnetism in unmagnetized steel.

In spite of his priority in the discovery of the magnetic effect of the current Oersted, who after all was led to it by an utterly wild theory, did little. In the hands of one of the great experimental geniuses of all time, André Marie Ampère, however, this new plaything of science led to results of the greatest import. From September 18 to November 2, 1820, Ampère was able to present at every sitting of the French Academy new results of the greatest importance. First he showed that if the current were to flow from the positive pole of the battery to the negative, in the sense of Franklin's theory for static electricity, from the feet of the observer to the head, and if the observer were to regard the needle of a compass, he would find that the north pole of the compass needle was always diverted by the current to his left. This rule he observed to be universal and it is to-day stated more conveniently as the right-hand rule as given on page 84. The rule is sometimes erroneously attributed to Oersted. Ampère next showed the attraction between parallel currents in the same direction and the repulsion between oppositely directed parallel currents. He next constructed the helical spiral and showed that with a current through it, the spiral acted just like a magnet both as regards other magnets and as regards its orientation in the earth's magnetic field. He developed the astatic needle for detection of weak currents (see page 124). He also postulated that the cause of magnetism might be found in the existence of electrical currents in the molecules, a prediction which modern atomic discoveries has substantiated in an unexpected fashion (see page 225). Ampère also suggested the use of the magnetic effect of a current for telegraphy.

In the meantime, 1820, Biot and Savart discovered, by experiment and mathematical analysis, their famous law concerning the force produced by an infinitely long straight conductor carrying a current (page 87). They found that the resultant force of such a conductor acts on a magnetic pole in a plane normal to the conductor and at right angles to the perpendicular on the conductor from the

pole, and that the force is inversely proportional to the distance between pole and conductor. In 1822 Schmidt showed that this law implied that the force between any point of the conductor and pole varies inversely as the square of the distance.

In 1821 Faraday using a new powerful battery showed by means of the Faraday disc (page 217) that a magnetic field and a current could so interact as to produce continuous motion. This machine which Faraday recognized as impractical for application was, however, the forerunner of our great electrical age. The Faraday experiment was extended in its application by Ampère, who caused a current to make a magnet rotate continuously. Improvements in batteries in 1822 led De la Rive with 380 cells and Davy with 2000 cells to discover the electrical carbon arc. This was the brightest artificial light source made by man, and by 1870 with the development of electrical generators was used for the first electrical lighting in streets.

Finally in the study of rotational motions caused by two circuits, instead of magnet and one circuit as theretofore, Ampère discovered his now famous *fundamental law of electrical currents.* This law for the first time placed the study of electrical currents on a quantitative basis, so that further advances could be made, and gave not only a definition of unit current on the Newtonian scheme, but also gave the basic law for the calculation of all interactions between currents and currents, or currents and magnetic fields. This fundamental law, which we shall term *Ampère's law* from now on, was first published in the *Annales de Chimie et de Physique* in 1822, and more completely in his Mémoire in 1823. It is strange that in modern texts so little attention is paid to the significance of the famous law, and that it is even rare to find the law coupled with the name of its great discoverer if mentioned at all, except in the definition of current. The law in its simplest form states that the force of a current i flowing through a short length ds of a conductor perpendicular to the line joining it to a magnetic pole of strength m placed r cm away from ds is given by.

$$f = \frac{idsm}{r^2}. \quad \text{(See page 85.)}$$

The original law was much more general and applied to the case of two small elements of conductors ds and ds' interacting on each other. r is the distance between them and θ and θ' are the angles between r and ds and r and ds', while ϵ is the angle between the lines ds and ds'. Then if we properly choose our units we can write that

the force f between the elements, for two currents i and i' in ds and ds', is

$$f = \frac{ii'dsds'}{r^2}(\cos \epsilon - \tfrac{3}{2} \cos \theta \cos \theta').$$

This law leads at once to the simpler law above if a magnetic pole be substituted for one of the currents i'. The simpler law at once leads to the two forms of the definition of unit current, so often quoted without mentioning the discoverer, which is given on pages 85 and 86. As soon as the law was established many devices came into being for the absolute measurement of electrical currents. The tangent galvanometer of Pouillet, 1837, was the first of these devices capable of absolute measurement which led to reliable results, and it is for this reason that the instrument is usually inflicted on elementary students of physics. The first galvanometer of the more modern type in which a movable magnet is suspended inside a fixed coil and which was read by use of mirror and scale was due to W. Weber in 1846. Thus it is seen that really further development of the measurements of currents belongs well in the great electrotechnical period.

The discovery of Ampère, however, led more immediately to other important discoveries concerning currents and the circuits in which they flowed. The earliest experiments of Ritter on galvanic cells had established the usefulness of metals for the conduction of the current as well as the heating effect of currents in wires. Davy had even suggested that metals had a "conductivity" which he assumed to vary inversely as the heating effect, and by 1822 he had arranged the metals Ag, Cu, Pb, Au, Zn, Sn, Pt, Pd, Fe in the order from best to poorest conductors, a rather crude result to say the least. It had also been noted that magnetic and other effects were strangely related to the number of galvanic cells in series, to the area of the plates of the cells, and to the nature and lengths and area of the conductors. It was, however, difficult to formulate clearly the laws concerning the conductivity of metals, or those concerned with the strength of currents, owing to difficulties encountered with the early batteries such as lack of constancy, polarization, and internal resistance. Around 1820 Schweigger and Poggendorf had independently devised an instrument called the multiplier for arbitrarily measuring and comparing electrical currents, using what is essentially a crude moving coil galvanometer. The current to be measured had its effect multiplied by using a number of turns of wire in a coil, and the interactions of coil and a magnet gave deflections whose measurement indicated current intensity. Currents could then be compared. Assuming that deflec-

tion and current in one of these instruments were proportional, G. S. Ohm began a study of the currents through conductors of different lengths and properties in a systematic fashion in 1825. He got around polarization of his cell by measuring only the initial deflection on closing the circuit. Ohm then evaluated the relative conductivities of a number of metals in a far more satisfactory fashion than Davy had. At the same time Barlow similarly investigated conductivity and found that the effect of changes in length produced changes in conductivity inversely proportional to the changes produced by changes in area of cross-section of conductors of the same substance. He also showed that the current down the whole length of a long uniform wire was constant. Ohm verified these results and improved on them. He then resorted to thermoelectric sources of potential from thermal junctions at fixed temperature differences (making use of the newly discovered Seebeck effect, see page 166), and therefore avoided further battery troubles. He at once discovered that the strength X of the magnetic action of the current varied as $X = \frac{a}{b + x}$, where a and b were constants, and x the length of the conductor used. He next studied the effect of using more than one battery (i.e., in this case thermo elements) both in series and in parallel. The results were expressed in the simple equations for series and parallel connections with m cells shown below:

$$X = \frac{ma}{mb + x} \quad \text{and} \quad X = \frac{a}{\frac{b}{m} + x},$$

and given on page 162. He also observed that increasing the temperature of a metal conductor effectively increased the resistance. As a result of his researches Ohm in 1827 published his book, "The Galvanic Chain." The organization of his results into a comprehensive theory was made possible to Ohm through the publication in 1822 by Fourier of his mathematical theory of heat conduction. The analogy of the experimental equations of electrical currents in conductors with characteristic equations of heat flow in solids led Ohm very quickly to write his law of conduction as

$$e = kA\frac{\partial U}{\partial N},$$

where k is the constant of proportionality depending on the units and nature of the substance, A is the area of cross-section of a conductor

$\frac{\partial U}{\partial N}$ is the rate of change of "electrical density" along the wire, and e is the effectiveness of the current. If we write instead of ∂U the change PD of "electrical density" over a length $\partial N = l$ of conductor studied we have writing i the current for e, to which it is proportional, $i = k\frac{A}{l}PD$. But $k\frac{A}{l} = 1/\frac{l}{kA} = \frac{1}{R}$ where R is the *resistance* of the circuit, and we have Ohm's law in its familiar form $i = \frac{PD}{R}$. Here PD and R are new concepts in the phenomena of current electricity, which really are analogies borrowed by Ohm from the resistance to the flow of heat in solids. The term PD in electricity is analogous to the temperature difference causing the flow of heat. It represented then a sort of difference of electrical pressure which was responsible for driving the current through the circuit. Its significance from an electromechanical viewpoint was already *capable* of interpretation as a result of the development of static electricity, and is fully treated in Chapters VI and XV. It is doubtful if even in statics the electromechanical concept was known until the time of Gauss in 1840. Hence it is most doubtful that Ohm himself recognized the relationship in any form other than the temperature analogy given. The term R represents the resistance to the flow of electricity offered by the substance. It is the reciprocal of its conductivity which again was defined in analogy to heat conductivity. In the later form Ohm's law then states that the current produced is merely the potential difference driving the current divided by the quantity which we will call the resistance of the circuit. In his researches Ohm recognized that in a long uniform wire with a current i in it there was a *continuous fall of the potential* or electrical pressure and that the fall PD over any length of resistance R through which i flowed was given by the relation $iR = PD$. The concept is one of the most important concepts which we have in dealing with current phenomena. In 1831 Fechner carefully verified Ohm's law in a large number of cases, and in 1837 owing to the improvement in cells the law was definitely established by Pouillet. It was the latter's work which first reached England and France, and for some time these countries disregarded the fact that Ohm was in reality the discoverer of the law which now bears his name. The actual experimental test of the existence of the fall of potential down a conductor and the proof that $PD = iR$ was due to Kohlrausch, who in 1848 actually measured the potential with an electrometer. The further great developments of the law were due

to Wheatstone in 1843 from the standpoint of an experimental extension in measurements, and by Kirchhoff in 1846 in the great systematic period. Kirchhoff's generalization of the law extended its use to all types of circuits and are extensively discussed in Chapter VIII. The further study of the concepts of resistance and potential in their relation to current really belong in the next period, the electrotechnical period. They were the result of the development of the first law of thermodynamics for which the foundation was firmly established by the experiments of Joule in 1843 on the mechanical equivalent of heat. In fact Joule used the heating effect of a current observed by Ritter and by Davy in his measurement of resistance. Joule showed that the heat was given by $H = ki^2Rt$, where k is a constant depending on units, i the current, R the resistance and t the time. This incorporated the electromagnetic phenomena even more closely into the Newtonian mechanical scheme and gave a new method for measuring potential difference experimentally. For since $iR = PD$ the equation of Joule becomes $H = kiPDt$, and as i, H and t could be measured and k was known PD was measured directly.

As is seen these discoveries and their further developments continue in an unbroken fashion well beyond the date 1831 which we have chosen to mark the beginning of a new period in the history of electricity. In fact 1831 does not mark such an abrupt change in focus of interest as did the discovery of galvanic currents. What happened was that a new and striking discovery added impetus to a natural trend in the development of the subject by indicating in a striking way the future possibilities of electricity, thus leading to a perhaps more rapid development of what would have inevitably followed.

c. The Electrotechnical Period, 1831–1865.—The period, while chiefly concerned with a logical development of the new quantitative facts concerning galvanic currents, begins with Faraday's discovery of electromagnetic induction, and terminates perhaps somewhat arbitrarily with the great period of systematization of electrical knowledge whose proper development is characterized by the publication of his electromagnetic theory in 1865 by James Clerk Maxwell.

The marked magnetic effects of an electrical current had long convinced Faraday that some reciprocal relation existed between electricity and magnetism, and that as electrical currents produced magnetic effects, magnetism should produce electrical currents. He was so firmly convinced of this that time and again he tried to detect the relationship, each time failing, and strange to say as we regard his unsuccessful experiments we are amazed to note that he came so inexpressibly near to succeeding that the failures are almost inexplicable.

Meanwhile Arago had made an observation which baffled the early physicists. In 1824 he observed that a magnet swinging about its pivot directly over a metal plate was damped in its vibration very much more rapidly than when swinging freely or over an insulator. He also found that if the metal disc under a magnet needle were rotated the needle rotated with it. All attempts at an explanation had failed, but the experiment continued to intrigue Faraday, who believed he saw in it the sort of action he was expecting to find. In December, 1824, he performed an experiment in which a magnet was made to pass into a conducting helical coil. Had the apparatus attached to the coil been simpler he would have observed his much hoped for effect. Complexities made this detection impossible. In November, 1825, he passed a current through a wire lying near another wire connected to a galvanometer. Possibly owing to the placement of his key for closing the circuit and possibly because of the feeble effects produced he did not notice the momentary change produced on making and breaking the battery circuit. On December 2, 1825, and again on April 22, 1826, he made experiments which gave no result. On August 29, 1831, he again began his experiments and in his note book there is a notation "Experiments on the production of electricity from magentism." The first experiment records the successful achievement of his aim. He used an iron ring 6 inches in diameter of $\frac{7}{8}$ inch soft iron, with two coils, one *A* of many turns of wire (3 coils of 24 feet in length) the other *B* of 2 coils of 30 feet in length. The separate coils could be connected in series or used separately. He used a battery of 10 plates 4 inches square which could be attached to *A*. *B* was connected to the two ends of a long copper wire which 3 feet away passed over a magnet needle. When contact was made between the battery and *A* the needle gave a kick and when its contact was broken the needle gave a kick in the opposite direction. In ten days of experimentation over an interval of several months, from August 29 to November 4, he completed his researches and published his famous memoir delivered on November 24 before the Royal Society. The time between the ten days was presumably spent in the construction of the apparatus and preparations. He next showed that if an iron bar with a winding about it connected to what was virtually a galvanometer was suddenly thrust between the poles of a magnet there was a deflection in one sense while the bar moved. On removing the bar there was a deflection in the reverse sense. He then showed that a bar magnet thrust into a helix produced the effect. Next he made a copper disc turn between the poles of an electromagnet and connecting a galvanometer to its edge and its axis obtained a current as long

as the disc moved. This is the famous Faraday disc experiment discussed on page 246, and is actually the first dynamo. Finally he showed that a copper wire moved relative to the poles of a magnet or a conductor carrying a current produced the effect. These experiments in their entirety completely proved the existence of electromagnetic induction. They further showed the salient laws concerned in it which he correctly generalized. The experiment with the Faraday disc proved to him the nature of the damping and motion observed by Arago in 1824. He was at once able from his laws of induction and these experiments together with the motor rules to show that Arago's experiments were the results to be expected from induction. The discoveries mentioned above lie at the basis of all the great electrotechnical achievements of the present day, for they enabled, through the improvement of dynamos the efficient conversion of steam motive power into electrical energy. In the earth inductor, page 243, Faraday made the first approach to an efficient dynamo in 1832. The first practical dynamo should however be credited to Werner Siemens in about 1866, less than thirty-five years after the discovery of the fundamental law.

The further developments of Faraday's induction experiments were done largely by Lenz. In 1835 Lenz showed that the intensity of the induced current was proportional to the number of turns in the coils, and derived the expression for the most effective winding taking account of the resistance of the wire. He next showed in 1839 that the strength of the current was in proportion to the magnetism created or destroyed inside the coil. Lenz generalized the laws of induction in the following way, namely, that the induced current always flows in such a way as to generate a magnetic field which will oppose the change in the magnetic field causing the induced current. This law was extended to not only magnetic fields of magnets but also magnetic fields due to currents in about 1851–1854 by Felici and by Gaugain. The extension came together with the recognition of the law of conservation of energy. For unless Lenz's law holds we could by induction effects get energy out of nothing. The more sophisticated statement of the laws of induction given above is perhaps the easiest form to remember and thus the most generally useful form of the law.

There was in the domain of induction, however, still a further important discovery to make. Jenkin in 1835 noticed that the spark on breaking an induction circuit was markedly strengthened if the wire were increased in length, especially if in the form of a coil. It remained for Faraday, however, to explain the effect, or at least to

observe the nature of the effect. If a circuit in a coil be left closed a parallel auxiliary circuit containing a galvanometer will show a current in the same sense as the current in the coil. By holding the galvanometer needle at the point where the steady current would put it, Faraday, having opened the battery circuit, observed that the deflection of the galvanometer was *momentarily* greater when the steady current was suddenly reestablished. The *instantaneous* current Faraday called the "extra" current. It was due to an E.M.F. generated in the coil opposing the making of the current in the coil, but by the galvanometer connections able to flow through the instrument in the same sense as the steady current. This phenomenon was, however, not clearly understood until Joseph Henry, an American high-school teacher working under the most adverse conditions, established the phenomenon as that which we would to-day term self-induction, a designation given by Henry in his explanation. Joseph Henry, independently and almost, but not quite simultaneously with Faraday discovered induction, so that it is not surprising that the later achievement should have been his. He had been constructing and studying powerful electromagnets and he, as had Jenkin, noted the heavy sparks obtained when he broke these magnetic circuits with heavy windings. He studied the effect of the different windings on the spark and found that the coil in the form of a close helix gave a very powerful spark but the same length of wire if straight gave little effect, and the less effect the shorter the wire. He concluded that the effect was the inductive effect of the change of current, and hence the magnetic field of the wire on itself. This was announced in 1832 in a paper on "Electrical Self-Induction in a Long Helical Wire." It is seen that the paper really antedates those on the same phenomenon by Faraday and Jenkin. However, communications with Europe were then none too good, and it took a long while for the world to recognize the merit of Henry's work. In 1842 Henry was led to the conclusion that the discharge of a condenser was under some conditions oscillatory. The fact was definitely proven experimentally by Feddersen in 1858, sixteen years later, and proved of great importance in the discovery of the Hertzian waves. With the discovery of self-induction and the discovery of the nature of dielectrics and the definition of electrical capacity, probably by Kelvin in 1850–1854 (see page 196), the development of the six important properties of electricity was complete. These fundamental properties are electrical quantity, electrical current, electrical potential, resistance, capacity, and self-induction. The stage was thus set for the later unification of electricity and magnetism into a whole system after the beginning by Gauss and Weber in 1839 and

ɜrminating in the work of the International Electrical Congress in 881.

During the electrotechnical period two more fundamental investiɜtions of an experimental nature which were completed, must be menoned. Both of these again involve basic investigations of Faraday. ollowing the development of the Leyden jar little if any progress as made in the study of the fundamental property of these jars of oring electricty. Coulomb's measurements gave rise to a new measrement of the quantity of electricity, and these measurements brought oulomb near to the definition of capacity. Certainly Coulomb was ware that two conducting spheres of different size when brought in ɔntact, after separation did not have the same quantity of electricity, ıe larger one having a larger charge in proportion to its radius. Couɔmb must also have had some notion of the degree or intensity of ectrification as shown by the divergence of the gold leaves of an elecoscope. That Coulomb was aware of the relation that $Q = CV$, here Q is the quantity, V the potential, or degree of electrification, nd C is the capacity of the system for holding electricity correspondıg to a definite degree of electrification, is doubtful. Certainly this ղuation does not appear in the continuation of Coulomb's researches y Faraday in the years 1837 to 1838. It is probable that the idea of ectrostatic potential had not been sufficiently developed by the ıathematical physicists even in Faraday's time to lead to this result. araday having finished his studies on electromagnetic induction, as ell as his work on conduction of electrolytes, turned his attention) what he called electrical induction. That is the effect of charges ı *non-conducting solid* and *liquid bodies*, which he termed *dielectrics*. is studies convinced him of the fact that charges appear in such ielectric bodies only in the inductive form, i.e., in a separation of ղual amounts of positive and negative electricity, in analogy to the duction in metallic conductors already clearly understood. He ıowed that a single sign of charge is never "soaked up or absorbed" y a dielectric. This we know to-day to be somewhat erroneous, for e can actually load a dielectric with electrons or remove them ıerefrom. For this, however, very much more drastic measures e needed than were employed in those days. To study the pheɔmenon better Faraday constructed the spherical condenser (page ɔ9) and measured the quantity of electricity stored at a given potenal (i.e., for a potential that would jump a spark gap of $\frac{1}{16}$ inch or ıereabouts). He measured the charge by means of a Coulomb toron balance. Here we see that Faraday clearly related the quantity ored to the potential and eliminated the latter as a variable by keep-

ing it constant. The charge on the condenser was measured for various substances placed between the concentric spheres of the condenser, such as gases, turpentine, naphtha, sulfur, glass, etc. In this way he discovered that the charge was greater for liquids and solids than for gases, or for vacuum between the spheres. The ratio of the value of the charge in these different cases relative to the charge in air or better in vacuum (the effect of air was nearly too small to measure) he called the *specific inductive capacity.* This ratio was always greater than unity, the value for empty space. The phenomenon Faraday explained correctly as due to some sort of electrical "polarization" of the particles of the dielectric. He thus discovered essentially the nature of electrical capacity and the effect of matter on this capacity. When and by whom the exact definition of capacity was achieved is difficult to ascertain. It was doubtless developed and the capacity of condensers calculated some time between these researches and the period around 1853 when the oscillatory discharge of a condenser was predicted from equations by Kelvin on the basis of the effect of capacity on circuit. The accurate concept of capacity must have come into use in its complete form as a result of Faraday's researches and the mathematical physical researches of Gauss, 1839, and of Green, the latter's work being extended by Sir William Thomson (Lord Kelvin) in 1850. Green, 1832, and later following Green, Kelvin, 1850–1854, established a clear conception of the electrostatic potential function, and it is stated by Hoppe that these papers established clearly the theory of the Leyden jar. The definition of the electrical and magnetic units in the fashion used in this text was due to G. F. Gauss, who developed the ideas of potential functions, etc., independently of Green and attempted to develop an absolute system of units. In magnetism using the milligram, the millimeter and the second Gauss defined the absolute unit of pole strength, and developed the methods of measurement of magnetic pole strengths on an absolute basis which are described in Chapter IV. The first experiments were performed by W. Weber in 1836. In 1839 Gauss arrived at a new and very important definition of electrical potential applicable in general form as well to magnetism and gravitation as to electricity. The relation $V = \sum \frac{q}{r}$ he called the *potential* of all the charges q (masses or poles), on a unit charge (mass or pole) at r cm from each of the separate charges, whose sum is indicated by the letter Σ. This is the formulation that is in a more elementary way given in the beginning of Chapter XV. He also relates the potential to the work required to move the body up to the point from infinity. The Gaussian theory was

applied by F. E. Neumann in 1845–1847 to the case of electromagnetic induction which by its means he formulated in an elegant manner. This formulation was in agreement with the principle of conservation of energy as shown by Helmholtz in his famous paper on conservation of energy in 1847.

Weber in 1840, carrying out experimental determinations required by the Gaussian formulation, made the first absolute measurement of electrical current using the earth's field as measured by the Gaussian method and the tangent galvanometer, and then determined the electrochemical equivalent of Faraday in absolute measure. By 1845 the laws of divided circuits had been worked out by Kirchhoff, then a student of Neumann, and were at once applied to Wheatstone's bridge so that the modern methods of resistance measurement were once and for all made possible. (See Chapters VIII and X.)

The final great experimental discovery of this epoch to be discussed here, although it antedated the one on dielectrices, was the quantitative study of electrolysis by Faraday. The earlier discoveries in this field by Ritter, Davy and others have already been mentioned. In 1834 Faraday began a study of the conduction of the electrical current in aqueous solutions and ended by finding his famous laws of electrolysis which can be summed up in one sentence in the modern terminology by saying that 96,500 coulombs of electricity will liberate 1 gram-atom equivalent of any substance from solution. (See Chapter XI.) The full significance of these laws was not recognized by Faraday, but Helmholtz in his famous Faraday lecture on Faraday's death in 1867 pointed out that, taken with the then theoretically established atomic theory, the laws of Faraday pointed definitely to the atomic nature of electricity, in contradistinction to fluid theories.

In 1847 Weber began to develop the laws of the interactions of currents in wires which led to the development of the absolute dynamometric measurement of current. Finally in 1852 Weber began his measurements of current, potential and resistance in absolute measure, using the Gaussian system. His work led to the independent definition of current in the electromagnetic system (page 80) and in the electrostatic system (page 183) and to a means of determining the ratio between them. He found that i in E.S.U. $= i$ in E.M.U. $\left(\frac{c}{4}\right)$ where c is a constant, for potential e (in E.M.U.) $= E$ (in E.S.U.) $\times \frac{c}{4}$, and R (in E.M.U.) $= R$ (in E.S.U.) $\times \frac{c^2}{16}$. c was found by measurement to be 4.394×10^{10} cm./sec. The velocity of light had been

found as about 3×10^{10} cm./sec, a very important coincidence as later results were to show, although the factor $\frac{1}{4}$ above constituted an error as Maxwell's later formulation was to show.

With these results one can fairly well see the nature of the period covered. While it is characterized and perhaps overshadowed by the discovery of electromagnetic induction its development implies more than just this. It was the period in which the last of the important discoveries necessary for the next great period were made. It is further characterized, not only by the applications of these discoveries to practical problems, but also by the great development in technique in measurement of these newly found concepts. With this there was also developing an excellent theoretical background of generalizations which ended in the early attempts of Weber and Gauss to establish absolute standards, and to correlate the separate phenomena into a whole. Hence by 1865 the stage was well arranged for the brilliant generalizations of Maxwell and their substantiation, which characterized the fourth great period and completed the Newtonian formulation of physics.

d. The Systematic Period, 1865–1896.—The fourth great period, called the systematic period, begins with Maxwell's publication of his famous paper on the electrodynamic theory of the electromagnetic field in 1865. It continues until the discovery of the electron by Sir J. J. Thomson in 1896, resulting from the study of the nature of the discharge causing x-rays. The period is named the systematic period, for it was the period in which the greatest advances were the theoretical ones in which not only all the facts of electricity and magnetism but most of the facts in physics were built into a great and consistent system based on Newton's laws.

As was seen in the last period the developments towards the end of this period were beginning to be dominated by advances of a theoretical, mathematical, physical character. As was seen attempts had been made by Weber, W. Thomson (Lord Kelvin), Gauss, Neumann and Kirchhoff to generalize the laws of electricity and magnetism. To this group must also be added the names of Helmholtz and Riemann. The attempts led towards the end to controversial questions until Maxwell appeared in the field. The concepts of Faraday which guided Faraday's fruitful experiments, and conceived of the electrical and magnetic forces as being represented by lines of force acting like self-repulsive elastic bands, had made a deep impression on Maxwell. The further analogies of Thomson between elastic and electrical phenomena also played their part. In his first paper in 1856 Maxwell clearly shows the influences of these ideas on his thinking, and his con-

tribution fundamentally lay in placing these vague qualitative concepts of Faraday in concise and accurate mathematical language. The results of Maxwell's studies are summarized in his "Treatise on Electricity" published in 1873. The details of these studies are far beyond the scope of this text and hence can hardly be given. From his definitions of the electrical and magnetic field and from the electromagnetic relationships Maxwell developed his famous equations on the electromagnetic theory. These equations led at once to the result that if a charged body be accelerated there travels out through space, which Maxwell assumed filled with an imponderable substance, ether, an electromagnetic pulse. If the motion of the charged body is simple periodic then there is emitted a train of electromagnetic waves. A study of these waves led him to the conclusion that they must travel with the velocity of light, a fact which was established by the experimentally measured ratios of electromagnetic and electrostatic units made by Weber, see page 39. His work went further and indicated that light was merely such an electromagnetic wave motion. He then showed the remarkable relation between the dielectric constant of transparent colorless substances and the index, (n), of refraction for light waves. The index of refraction is determined by the ratio of the velocity of light in vacuum and in the substance. According to Maxwell's theory the velocity of electromagnetic waves is given by $v = \frac{1}{\sqrt{\mu D}}$, where D is the dielectric constant and μ is the magnetic permeability, nearly 1 for most substances. Hence $v^2 = \frac{1}{\mu D}$, so that $n^2 = \mu D$, assuming the velocity of light in vacuum to have the value unity.

The theory of Maxwell received its most striking confirmation in 1887 in the hands of H. Hertz. Following the early speculation of Henry that the discharge of a condenser was oscillatory, Feddersen in 1858 actually succeeded in establishing the fact. The complete theory of the discharge was later worked out by W. Thomson in 1853 as part of his extensions of Green's theories in 1850–1854. The possibility of the use of such oscillatory sources of energy appealed to Hertz, who began investigating oscillatory discharges. By using tuned inductance coils for sending and receiving he succeeded in transmitting sparks across the length of the laboratory table. The intensity was increased when he used both capacities and inductance coils, (i.e., tuned heavily oscillating electrical circuits, with relatively little damping). Hertz succeeded in proving that these waves were reflected from metals and that they traveled with the velocity of light. Next he

reflected them in parabolic mirrors and refracted the waves as well as diffracting them.

In the same period came notable advances in measuring technique and the formulation of a consistent system of electrical units by the International Congress of 1881. In the development W. Thomson (Lord Kelvin) played a great rôle as did many others. It was also in this period that the world began to reap the reward of Faraday's discoveries through the development of electricity in driving cars, and lighting streets and houses, as well as in many other directions. The application of the laws of thermodynamics to electrical phenomena, especially to radiation, led to Kirchhoff's radiation laws, the Stefan-Boltzmann law and the Wien and Rayleigh laws of radiation. The latter laws were deduced on the assumption that electromagnetic waves and light waves were identical, and that light actually exerted a pressure. The failure of experiment and observation in the case of the radiation laws to agree, led to the enunciation of the quantum theory by Max Planck in 1900—a theory which was to play an important rôle in showing the limitations of Maxwell's great generalizations. The Michelson-Morley experiment was also performed at this period, and strange to say it was this apparently innocent experiment, intended to prove the existence of Maxwell's ether, that ended in disproving its existence in the sense in which it was intended, and ultimately led to the great theory of relativity of Einstein, 1905.

e. The Atomic Period, 1896–1919.—The fifth period, termed the atomic period, begins with the discovery of the electron, the negative unit of electricity, and dovetails so closely into the next or quantum period that dates are hard to fix. Perhaps 1919 is as good as any date to terminate this period.

As a result of a study of the bluish streamers observed in exhausted tubes when a spark giving out Roentgen's newly discovered x-rays passed through them, J. J. Thomson, 1896–1897, showed these streamers to be negative charges of electricity having a charge probably equal to that of the Cl^- ion in electrolysis, and a mass of about $\frac{1}{1800}$ of the H atom. This discovery was made independently nearly simultaneously by Lenard. The positive or canal rays had been observed by Goldstein in 1886. Shortly after Thomson's work the positive rays were extensively studied by M. Wien. By 1912 J. J. Thomson had shown them to be positively charged atoms of matter and had devised a method for measuring the ratio of their charge to their mass, and hence their atomic weight. The discovery of radioactivity by Becquerel in 1896 and the work of the Curies, Rutherford and many others led to the conclusion that the electrons and

positive helium atoms, the α rays, were building-stones of matter. The electromagnetic theory of matter was then quickly translated into the terminology of the electron by H. A. Lorentz, Richardson, J. J. Thomson and others so that the development of theory nearly kept pace with that of discovery. Nearly, but not quite, for ever since Planck's quantum law of radiation, theory began to find problems which it could not cope with. In 1908 Jean Perrin by a brilliant series of experiments based on theoretical deductions of Einstein and Smoluchowski showed the reality of the eternal heat motions of molecules and proved once and for all the molecular theory of matter, especially gases. This was a theory dating from Avogadro and Joule, and later developed by Clausius, Maxwell, Boltzmann and many others. By 1911 Millikan succeeded in measuring the *charge* on the electron quite accurately and bettering the value obtained almost simultaneously by Rutherford from the charge on the α particles.

Finally in this period through the photoelectric effect (Chapter XXVII), and the collisions between electrons and molecules or atoms, definite proofs of the quantum relationship of Planck were established. This meant that it appeared as if even radiant energy were emitted, absorbed, or even possibly moved in units or atoms. Truly this was an atomic period.

In this period as well Einstein in 1905 developed his special theory of relativity and in 1913 his general theory, while von Laue proved that x-rays were electromagnetic waves and measured their wave lengths. This was a very important forerunner of the next period, for the x-ray was to figure even more strikingly than light in the quantum theory. In this same period Rutherford, 1911–1913, developed the proof of the nuclear atom and Bohr (1913), showed that for the stability of such an atom the application of the quantum theory was essential, while classical electrodynamics was utterly powerless to predict the phenomena.

f. The Quantum Period, 1919–1927.—The sixth great period, the period of the quantum, can best be set as beginning in 1919, following the resumption of scientific work after the World War, although it really began in 1900. It terminated only in the period 1926–1927 with the development of the wave mechanics.

As stated, Planck showed in 1900 that the radiation laws can be correctly derived only if we assume that radiant energy is absorbed or emitted in units of value $h\nu$, where ν is the frequency of the emitted light, and h is a new universal constant of nature, 6.55×10^{-27} erg secs. This was followed by the discovery that in the well-known photoelectric effect (emission of electrons from substances by the

action of light) the velocity of the emitted electron did not depend on the intensity of the light but depended on its frequency according to the law $\frac{1}{2}mv^2 = h\nu - h\nu_0$, where $\frac{1}{2}mv^2$ is the kinetic energy of the electron, ν the frequency of the incident light, and ν_0 the frequency of light needed to just remove the electron from the surface. The photoelectric effect had been discovered by Hallwachs in 1888. The above law was theoretically deduced by Einstein in 1905, who later retracted the assumptions involved in the deduction. A limited and incomplete verification of law had been achieved by Ladenburg, Richardson, K. T. Compton, and Hughes. In 1915 Millikan succeeded in accurately proving the law in spite of Einstein's withdrawal of his deduction. With the study of x-rays the law was still further established. This law while discordant with the classical electrodynamical considerations was not nearly as bad as another feature of the photoelectric effect. Light, no matter how feeble, was found to be able to liberate electrons the instant it struck the metal or surface from which electrons could be liberated. However, the intensity of light falling on the most liberal space that could be allowed an atom could never in years suffice to give the energy $h\nu$ to the electron on a classical electromagnetic wave basis. It was therefore a great temptation for some to return to the old Newtonian view that light was corpuscular (i.e., moved in strings and struck in a limited area) even in the face of the remarkable success of the electromagnetic wave theory of light. In the application of the quantum laws to the radiation of light by the Rutherford atom, Bohr, Sommerfeld, and others achieved such notable success in not only accounting for atomic structure, but in quantitatively deducing the hundreds of spectra of atoms and molecules, that no one could doubt the quantum principles in this connection. Again in that almost marvelous year for discovery, 1913 (year of discovery of x-ray diffraction, the Bohr theory, the establishment of the Rutherford atom) there was observed perhaps the most direct proof of the quantum action. James Franck and G. Hertz shot electrons into atoms of Hg and He at various speeds. They observed that under some circumstances these were reflected from the atoms with *practically no loss of energy*. On classical electrodynamics an electrical charge shot at an atom composed of electrons and a central positive charge should lose energy at all impacts. The energy was observed to be lost only when the electron had an energy in excess of a critical value. In such an impact it lost not a fraction but just this critical energy. This was just what quantum theory and Bohr's theory postulated. It was later shown that this minimum energy at the lower values went to exciting the first lines of the spectral series

of the atoms struck, and was just the energy required to excite them. The work of Franck and Hertz was therefore a direct quantitative proof of the quantum theory. Finally in 1923, A. H. Compton made an interesting discovery. If x-rays of a penetrating nature (short wave length), are shot into an element of low atomic weight where the electrons are to all intents and purposes practically free (i.e., the energy to excite them is negligible compared to the energy $h\nu$ available in the beam), the rays reflected are shifted to a longer wave length λ or shorter ν by an amount dependent on the angle between the initial beam and the scattered beam. The law found was that $\lambda_\theta - \lambda_0 = \frac{h}{mc}(1 - \cos\theta)$, where λ_θ is the wave length observed at an angle θ, λ_0 is the initial wave length of the impacting x-rays, h the Planck constant, m the mass of the electron and c the velocity of light. This law was derived theoretically assuming the x-ray wave pulse to collide with the electron like a particle of energy $h\nu_0$ and to be reflected at an angle θ giving the electron an energy which it would have had had it been struck by a particle with this energy which rebounded at an angle θ and gave the electron a kick. The loss in energy of the x-ray pulse and hence the energy given the electron caused the shift in wave length of the x-ray. This was taken by Compton and some others as a positive proof of the *corpuscular nature of light.* It made most physicists who had a great faith in the electromagnetic theory of light lose complete confidence in all the foundations of the classical wave theory of light. It was in this period of doubt and depression that the struggles of the mathematical physicists in searching for a solution began to bear fruit in the form of a new approach to physical problems that ushered in the new era or period entitled the wave mechanics period.

g. Wave Mechanics Period, 1926–1927.—The many failures at an explanation of the apparent dualistic nature of light (wave and corpuscle) began to bear fruit first in an almost ignored note by Louis de Broglie, a young French mathematical physicist, in 1921, in which he ventured to suggest that if, instead of considering the electron as a sort of electrical particle, it were considered as some sort of a complex wave motion represented best by an equation, the quantum actions might be explained and the wave theory of light retained, the wave electron interacting with light waves to give the desired result.

In 1925 W. Heisenberg and M. Born began a series of studies in which they showed that if the condition of the electron and light wave were expressed mathematically by means of groups of infinite series, called matrices, they could be made in some way to interact with other

electrons and light waves in a manner which seemed to agree with observation. To study these complex functions whose physical meaning was never quite clear they had to further develop a special kind of algebra, the algebra of matrices. While this technique was being developed, Schroedinger had independently arrived at what was an exceedingly convenient mathematical mode of expression for, or formulation of, the wave-like functions of de Broglie. At once the mathematical physicists began intensive and active development of this mode of approach. Shortly in the skillful hands of Schroedinger, Heisenberg, Pauli, Dirac, Oppenheimer, Sommerfeld, and many others, the new method was developed, enlarged, and applied to almost all the troublesome quantum problems in their elementary forms and found to yield complete coordination of all the apparent discrepancies between wave theory of light and electron behavior. It was further found that the matrix approach of Heisenberg and Born was merely a somewhat different approach to the same solution. Through the uncertainty principle of Heisenberg still more subtle difficulties were cleared up so that theoretically at least by 1928 optimism as to the state of physical knowledge had again begun to reign. This optimism, however, was brought on by something more than mathematical and theoretical success.

The radical assumptions as to the wave-like nature of electrons was difficult for the experimentalist, who must perforce live in a world of real mechanical pictures, to conceive. In 1927–1928 Davisson and Germer, who had for some years been trying to get a clear picture of the reflection of electron beams from crystals, succeeded in finding peculiar scattering patterns for electrons reflected from single crystals of Ni heated to remove adsorbed gas films. To explain these patterns they applied the wave mechanics expressions to their electron beams and found approximate agreement. Electrons appeared to be scattered from the Ni crystal not like bullets but like complex wave motions. In 1928 G. P. Thomson decided that as x-rays which are wave pulses are scattered into Laue patterns on passing through thin metal foils, electrons on the wave-mechanics picture should do the same. On shooting a beam of electrons through thin foils of mica and metals he observed instead of a single diffuse spot a perfect Laue pattern. Rupp also succeeded in obtaining these patterns and the agreement between theory and observation using wave mechanics was found to be fairly satisfactory.

Finally, in 1928–1929, O. Stern conceived of an even more striking proof. He shot beams of He atoms and H_2 molecules, not *into* a crystal nor *through thin foils*, but onto the simple crossed grating-like atomic

lattice at the surface of NaCl and LiF crystals. The scattered atoms and molecules formed exceedingly complex fan-shaped patterns. It was found that if the wave-mechanics picture were applied to the atoms of He and molecules of H_2 a perfect agreement good to two or three per cent with their observations was obtained in all cases. This striking proof leaves little to be desired, first in establishing the correctness of the wave-mechanics representation, and secondly in showing that the atoms and molecules of classical physics assumed to be like hard elastic billiard balls had to be replaced by mere equations of no clearly defined mechanical significance. What the future holds in this regard is not quite clear, but we are confident that the facts found thus far are correct and that we must be cautious in the future in not going too far beyond the facts in building our pictures of nature. Perhaps even it is safest to content ourselves with mere mathematical formulations, taking care not to exceed the scope to which these are experimentally verified, and leave the idea of Newtonian mechanical models only as aids in further experimental work. At any rate an interesting and stimulating, even exciting, future lies before those who venture into physics to-day, with a new clean slate to write on, and it should be a privilege to be undertaking the studies allied to physics in this period with all the future which it offers.

CHAPTER II

THE PHENOMENA OF MAGNETISM

3. THE DISCOVERY OF MAGNETISM

MAGNETISM became known to the world through discovery of natural magnets or lodestones consisting of pieces of the magnetic oxide of iron, Fe_3O_4, which occur in nature in the magnetized state. Whether they become magnetized through the action of the earth's field or possibly through the currents due to lighting discharges is not known, though the former origin is most likely. They are, however, frequently found in deposits of this ore. The property of lodestones of attracting pieces of iron was known to man in antiquity. It is likely that this attractive property was independently discovered wherever primitive man smelted iron, in the neighborhood of deposits of Fe_3O_4.

The art of smelting iron ores is mentioned in Greek legend first in the name of a sort of cult of roving iron miners and smelters, the Dactyls, who initially came from Phrygia inland of Troy which borders on the northeast corner of the Aegean Sea. These migrated both to Crete through Asia Minor and north and east to the islands Lemmos, Imbros and Samothrace in the Thracian Sea. Next came the Cabiri, a second more skilled group of iron workers who made of their trade a sort of cult. They spread north and east to Macedonia, and south and west to what is now Syria.

Iron reached Greece from Asia Minor. It was known in Egypt in the old kingdom 2900-2400 B.C. In Babylonia the dates are not known, but documents make it possible that iron in Egypt came initially from Babylonia or Phoenicia as there is no iron in Egypt. The dates for the movement of iron from Asia Minor and Phrygia to Greece are roughly as follows. In 1500-1400 B.C. iron was definitely in use in the regions mentioned above and perhaps earlier in Babylonia. By 1300 B.C. it was in Crete, and it appears to have reached Greece in quantity between 1200 and 1100 B.C. coincident with the Dorian invasion from the Danube countries which terminated in forming the Spartan state. That a knowledge of magnetism reached Greece at this time if not before is certain, for in the Phrygian iron mines about

Mount Ida, on the isle of Elba, Crete and Samothrace, there occurred with the regular iron ores (Fe_2O_3) that were smelted, deposits of what the Greeks called siderites or ironstone, Fe_3O_4, or magnetite, the natural magnetic oxide of iron. Hence early in Greek legend and in religious cults there appeared the Samothracian rings, rings of iron magnetized by contact with the magnetite. These attracted each other and occupied an important place in the early religious cults, such as those of the Dactyls and the Cabiri. These are clearly discussed by Plato (429-348 B.C.) who speculated on them on philosophical grounds. Earlier mention is made of magnetism in the works quoted from Thales of Miletus in Asia Minor (585 B.C.) who had traveled widely. Thales also mentions electrification of amber. The name magnet applied to these magnetized objects later came into use, and the name is attributed by some to a legend quoted by Pliny in which a Greek shepherd Magnes found a stone that attracted his iron-shod staff. Lucretius, 95 B.C., refers to the name magnet as being derived from the origin of magnets in the province of Magnesia lying in Thessaly on the sea coast between Mounts Ossa and Pelion. There is however no iron in that region. It is probable that the Magnetes of Thessaly, who because of overpopulation migrated across the Aegean from Magnesia, founded a city, now lost, called Magnesia in Ionia. They were later driven north and founded another city of Magnesia near Mount Sipylus in Lydia. It is assumed that it was the large deposits of magnetic oxides of iron in this region, which must have been exported to Greece, which furnished the source whence the name was derived. The city was destroyed by an earthquake in the time of Tiberius. The date of the supposed Magnesian migrations to Asia Minor lies between 700 and 1000 B.C. It is, however, not certain whether the migrations described above which were given by Pliny really took place. The probability, however, is great that the first commonly known magnets in Greece came from a town called Magnesia probably in Asia Minor close to a large bed of magnetite where iron was also smelted.

4. THE DISCOVERY OF THE COMPASS

The use of the knowledge of the directive force of the earth's magnetic field by means of compasses was not clearly recognized until a much later date than was the case with the attractive force. While Gilbert was the first to clearly distinguish the attractive properties and orientation of magnets in the earth's field, the knowledge of the orientation had been used much earlier by man in navigation. Frequent

attempts have been made to attribute this discovery to the Chinese in antiquity. Practically all such claims must be discounted owing to the very flowery and inaccurate language of the Chinese in describing what is supposed to have been the first compass. This supposed first compass was referred to as the "South-Seeking Chariot". Some claims that these were compasses of pivoted magnets in chariots used in land voyages before our era are definitely shown to be erroneous. The definite knowledge of magnetism in China is as follows: Iron was known in the Shensi province in 220 B.C. Records show imposts on iron as early as 685 B.C. Fe_3O_4 was present in the Shensi iron mines. Magnets were first definitely described in China in A.D. 121. By A.D. 324 it is more or less certain that the directive action was known and it is possible that magnets may have been applied to navigation in about this period. The Chinese admit that before A.D. 400 the present type of south-pointing chariot using a pivoted magnet concealed in a figure mounted on a chariot was not known. What the earlier term "South-Seeking Chariot" indicated is therefore merely a conjecture. In sea navigation the Chinese were very far behind other countries. Their first sea-going ships date from 139 B.C. and it is doubtful if the compass was used in sea navigation much before A.D. 1000-1100 when as in Europe we find definite proof of the use of the compass. That, however, the directive action of magnets was known in China in A.D. 700-900 is indubitable, for at that time we find measurements of the declination, or magnetic variation, of the compass. Hence we may conclude that while magnetic attraction was known in China perhaps centuries later than in Europe the directive action of the earth's field on a magnet was known perhaps as early as A.D. 400 and accurately studied by A.D. 900. Hence use of the compass in land navigation presumably dates from about A.D. 400-900. Its use in sea navigation, however, appears contemporaneously with the European use in A.D. 1000-1100.

The directive action of the magnet was not known in Europe until it suddenly appeared in completed form and extensive use in what is to all intents and purposes the modern type of compass. How the compass suddenly completed reached Medieval Europe is only a most interesting conjecture. That it reached seafaring peoples of all nations almost simultaneously at the period mentioned (A.D. 1000) is sufficiently witnessed by the rival claims for its discovery by most European countries. The earliest definite records of its use appear in the laws of Wisbuy, an early seaport of the island of Gottland in the Baltic. Here drastic laws appear for the punishment of seamen caught tampering with the compass, the punishment being to have the man

pinned to the mast by a knife driven through his right hand, from which he could be freed only by tearing his hand loose. These laws date from about A.D. 1000. This port was frequented by Goths, Russians, Swedes, Angles, Scots, Flemings, Saxons, Spaniards and Finns. The Finns were a people of Mongolian extraction and it is suspected that they carried the idea of the compass from China. The earliest compass card in Europe, which is unfortunately undated, is a Finnish card whose markings indicate that it was made for a latitude of 49° 20′ N, which would be the regions over which the Finns are thought to have traveled in coming from their Mongolian haunts. Thus we may fairly well assume that the compass as developed for sea navigation must have come from China through the migrations of the Finns, in a period perhaps lying between A.D. 600 and 1000, whence it reached European people through Wisbuy in from A.D. 1000 to 1250.

5. THE QUALITATIVE PHENOMENA OF MAGNETISM

As stated in the historical introduction the scientific discoveries in magnetism really date from the first investigations of Peter Peregrinus in A.D. 1256 and more properly from Sir William Gilbert, physician to Queen Elizabeth, as set forth in his book "De Magnete" in 1600. The facts concerning magnetism given below except for those dealing with electrical currents and those dealing with magnetostriction are essentially the qualitative facts established by these early investigators. It must be noted right here that magnetism evidenced itself only by its manifestation of forces (i.e., attractions, repulsions or orientations). It is clear therefore that until a quantitative analysis of forces was possible (i.e., an analysis only possible after the time of Newton, 1700) little progress in magnetism beyond that detailed below was possible.

If a magnetized piece of steel (a magnet) be pivoted at its center about a vertical axis it is observed that the magnet if left to itself always comes to rest with the same end pointing northward. This end of the magnet is thus endowed with a property which always makes it seek approximately the geographical north. It will also be noted that the opposite end of the magnet is endowed with a property which always makes it seek the approximate geographical south. Thus a magnetized piece of steel appears endowed at its ends by some power which makes the ends orient themselves in opposite directions which are related in a rough manner to the geographical poles of the earth. Any body which exhibits a tendency to orient itself in this fashion over the greater portions of the earth's surface, in the absence

of other magnets or of iron, may be considered as exhibiting *magnetic* properties and will be called a *magnet.** It is thus not unnatural to call the north-seeking end a *north magnetic pole* and the south-seeking end a *south magnetic pole.* Such a magnetic pole shows the power of attracting iron and unmagnetized steel in its immediate neighborhood, and this property, perhaps the first magnetic property noted, may also be used as a further criterion of magnetism.

If now two magnets be tested out so that their north-seeking or north magnetic poles can be marked it will be found if one magnet be placed on a pivot so that it is free to turn, that when the north poles of the magnets are brought into proximity of each other they repel one another with considerable force. Likewise a south magnetic pole will be repelled by a south magnetic pole, while a north magnetic pole will strongly attract a south magnetic pole and vice versa. Interchange of the two magnets will show that the actions of attraction and repulsion are mutual and reciprocal. Each magnet pole repelling and being repelled, or attracting and being attracted, as the case may be. These observations can be summed up in the terse statement that like magnetic poles repel and unlike magnetic poles attract, the first experiment having showed that there are two kinds or types or poles, or magnetism, termed north and south, a fact confirmed by the behavior exhibited in the second experiment.

Both Peregrinus and Gilbert had spheres turned out of magnetite which was magnetized. A study of the magnetism of these spheres by means of small magnets mounted on pivots showed that these spheres acted as if they had a bar magnet running through them in one direction. At one point on the sphere a north pole was repelled; at the other end of a diameter of the sphere, starting at the first pole, the south pole was repelled. At various points over the surface of the spheres the small magnets on pivots oriented themselves in definite positions so that their axes appeared to lie on surface lines over the sphere running from one of the ends of the diameter to the other. The analogy of this behavior to the behavior of magnets pivoted on the earth's surface was so striking that both observers decided that the earth's magnetic action on the pivoted magnets was due to north and south magnetic poles in the earth. They then considered that the earth itself must be a large magnet with one pole in the north geographical regions and another pole in the south geographical regions.

* Exception might be made to a helical coil of wire carrying a current which also exhibits this property. However the exception need not be made, for such a coil acts in other respects as a magnet and while strictly not a magnet certainly behaves like one.

Since a north-seeking magnet pole points northward, *the earth's magnetic pole in the north* must have *south magnetic polarity*, and *the earth's magnetic pole in the south* must have *north magnetic polarity*. Hence one can say that the earth's north geographical pole appears to have south magnetic polarity and vice versa. Actually, as we shall see, the earth's magnetic poles are oriented in this general fashion, except that the earth's south magnetic pole in the northern hemisphere is not exactly at the earth's north geographical pole, and the earth's north magnetic pole does not correspond in position exactly with the earth's south geographical pole (in other words, that the poles about which the earth rotates do not coincide in position with the earth's magnetic poles).

The use of the pivoted horizontal magnet for finding directions on the surface of the earth, through the use of the magnetic compass in navigation, has given the name of *magnetic compass* to the device used above, a name which will be used hereafter. As was seen, compasses are very useful for studying conditions produced by magnetization and furnish the first method of analysis if a body is suspected of being magnetized. The compass is used by jewelers continually to test for the magnetization of watches which operate irregularly. The action of a compass in studying the region about a magnetized body can be easily supplemented by the use of iron filings. These scattered over the region align themselves like compass needles and so enable one at a single time to map out a whole magnetic region.

The attractive and repulsive forces of magnets on other magnets led Gilbert to perform another experiment. He floated a compass needle or a magnet on a cork on water in the absence of other magnets. This floated magnet never *moved as a whole* to the north or the south, although if its axis was not parallel to the north and south magnetic line it *rotated* about its axis so as to have its north pole pointing north and its south pole pointing south. This clearly indicated to Gilbert that the earth's magnetism exerted a directive force on a magnet, as differentiated from the attractive or repulsive force of another magnet in proximity to the floating magnet. The more modern implication of this discovery is that owing to some uniformity of the earth's magnetism, or as we term it today the earth's magnetic field, the magnet is not displaced north or south, as its north and south poles have equal amounts of magnetism. Thus we can conclude that all magnets have equal amounts of north and south magnetism. That attractions and repulsions occur for magnet poles brought near the floating magnet is due to the fact that the action of the one pole of the magnet brought up is much greater on the one pole of the floating

magnet than on the opposite pole because its distance from the first pole is small compared to its distance from the opposite pole. It will later be seen that this signifies that the force between two poles varies very rapidly with the distance. We can then conclude that the magnetism of the two poles of a magnet is opposite in polarity and equal in intensity.

If a magnet be broken the poles at the ends are separated. However at each of the broken ends new poles of opposite sign appear. This with the other observations of magnetism establishes the fact that magnetic poles always appear in opposite pairs and that there is no such thing as an *isolated pole.* The only way in which one can study a pole isolated from its opposite companion is to use a magnet so long that the second pole is ineffective in its action owing to the distance. Thus in some experiments magnets of magnetized steel tape a meter long are used. The fact that a magnet which is broken gives two new complete magnets no matter how often it is broken early led to the idea that the magnetism must be inherently lodged in the ultimate particles (the molecules or atoms of the iron).

It was later found that magnetism is not confined to iron and steel alone, but that the elements nickel and cobalt show it nearly as strongly as does iron. In addition certain alloys of copper, tin and manganese known as the Heussler alloys show marked magnetic properties. The salts of iron, cobalt and nickel show feeble magnetic properties which will be discussed at another point. Outside of this the magnetic behavior of most bodies is feeble in the extreme, so that today we call the behavior ferromagnetism in contradistinction to the other weaker magnetic phenomena.

Magnetic forces will penetrate all non-magnetic substances such as wood, ebonite, brass, glass, etc. The forces are however much attenuated in passing through thin sheets of iron, unmagnetized steel, or other magnetic substances. This property is often made use of to screen out magnetic effects. In certain sensitive galvanometers it requires some seven or eight sets of soft iron screens each made out of three or four sheets of soft iron 1 mm thick to cut out the earth's weak field fairly completely. The cause for the screening action of iron is easily seen when we study another feature of magnetism observed by Gilbert. If a piece of soft iron be brought near a magnet it becomes a temporary magnet with its north pole opposite the activating magnet's south pole or vice versa. This is what is called induced magnetism. The screening action of iron can at once be explained by the fact that a continuous strip of iron going from the north pole of a magnet around to the south pole becomes an induced magnet which acts to

neutralize the magnetic poles by its opposite induced poles nearby. Another explanation which could be given for this will be seen to lie in the fact that such a piece of iron acts as a magnetic conductor, see page 229.

The action of soft iron filings in adhering to a magnet, like a mass of whiskers, is due to the inductive action of the strong magnet through a whole chain of filings which cohere due to the attractive forces.

In distinction to the action of soft iron, steel shows another type of action. If a magnet be brought near a piece of unmagnetized steel the steel shows the same inductive action as was exhibited by the soft iron. On withdrawing the magnet the steel will, however, unlike the soft iron, *retain* at least a portion of its induced magnetism; that is, it has become a more or less permanent magnet. By more drastic treatment with a pole of opposite polarity than the first pole it can be caused to reverse its magnetism and again be attracted. Soft iron is always attracted and loses most of its magnetism as fast as the inducing magnet is removed. The difference in action of the two lies merely in the fact that *once magnetized, steel retains its magnetism* to a high degree while soft iron does so only feebly. All degrees of this property known as *retentivity* are encountered in different samples, in general the harder steel being the more retentive. The difference depends on the crystal form of the substance.

Gilbert found that if a magnet be heated its magnetism is weakened or destroyed. The warmer a magnet the more rapidly it loses its magnetism. All magnets heated above a certain temperature (about bright red heat) cease to be magnetic. Jarring, shaking, or violent mechanical treatment causes magnets to lose their magnetism. Even the best steel magnets eventually weaken, a fact which is attributed to a very slow action of heat together with what is called demagnetization by the field discussed in Chapter XIX. To prevent this demagnetization these magnets are provided with a soft iron keeper, or two bar magnets are kept together in pairs in a box with opposite poles adjacent and soft iron keepers between, Fig. 9.

Magnets can be made by the following procedures, all except the last one having been observed by Gilbert.

(*a*) By the action of a magnet or a lodestone on steel, through contact or by stroking the steel with the magnet poles.

(*b*) By pounding a piece of iron or steel held in the direction of the earth's magnetic field. It is by the mechanical jarring in the earth's field that tools become magnetized, according to Gilbert.

(*c*) By heating steel or iron in a magnetic field and letting it cool.

(*d*) By the action of an electrical current in a solenoid on steel or

iron. This was discovered by Gay-Lussac and Arago in 1820 shortly after Oersted observed the magnetic effect of a current.

It was early observed in magnetizing a piece of steel that the amount of magnetizing treatment increased the strength of the magnet at first rather rapidly, and then more slowly. In all cases a state was finally reached beyond which the strength of the magnet could not be increased. In this case the magnet was said to be *saturated.* The physical aspects of saturation will become clearer when the magnetic properties of materials are discussed. The interpretation of the meaning of the phenomenon in its elementary form came however much earlier

It was also later observed that bodies being magnetized undergo a series of changes of length in the process. If a bar of soft iron be placed in a solenoid and a current turned on the bar will appear to be elongated if the currents are weak enough so as not to produce saturation. The elongation can be shown by fastening the lower end of the bar rigidly and leaving the upper end of the bar free to move so that in moving it rotates a small mirror by means of a lever multiplying device. As the current is increased and the current is switched on and off the movement of the spot of light from the mirror indicating a lengthening of the bar decreases in amplitude, becomes zero and reverses, indicating a contraction of the iron bar as fields producing saturation are reached. Hence weak magnetization causes an elongation while saturation causes a contraction. The phenomenon is called magnetostriction.

To get a better picture of the conditions surrounding a magnet the use of iron filings representing a mass of small compass needles about a magnet can be resorted to. If a piece of Celophane or other transparent substance be placed over a magnet lying in a beam of light and if light iron filings be dusted over the Celophane and the latter jarred the filings will arrange themselves in a set of regular patterns which can be projected on a screen. The patterns indicate the setting which compass needles would take in the regions about the magnet. It will be seen that the whole region around the magnet appears to exert a directive action on the filings. Such a directive action implies the existence of magnetic forces at all points in the space about the magnet. Hence we can say that the region about a magnet represents a *field* of magnetic force, the direction of this field in each place being the direction represented by the long axis of the small piece of filing. The soft iron filings are temporary magnets or compass needles due to the inductive action of the field. An idealized picture of the field about an isolated magnet is shown in Fig. 1.

It is instructive to observe such filing patterns as displayed by different arrangements of poles and magnets. The attractive forces between N and S poles are clearly illustrated by the continuity of lines of filings running from the N pole to the S pole, see Fig. 2. The

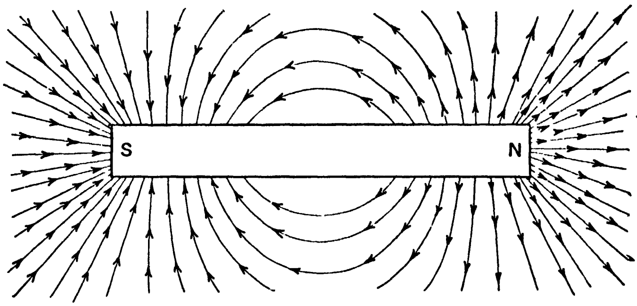


FIG. 1.—Idealized Field about an Isolated Magnet.

repulsive forces between the two N or two S poles can be seen by the recurving of the filing lines, leaving one pole as they approach the other similar pole, leaving a *neutral point* between the two sets of recurving lines as seen in Fig. 3. The neutral points of a magnet placed in the earth's field are shown in Fig. 4.

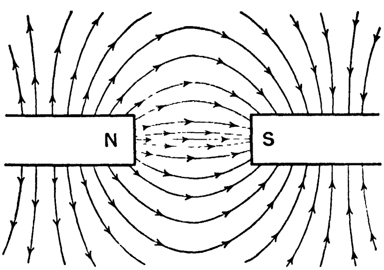


FIG. 2.—Idealized Field between Unlike Poles.

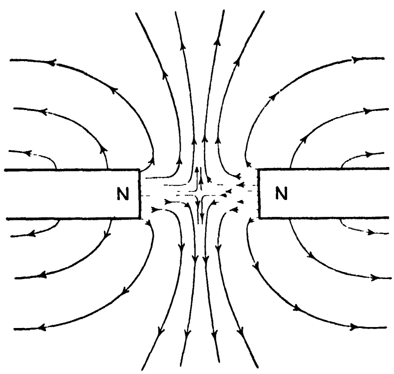


FIG. 3.—Idealized Field between Like Poles.

A more important observation is the fact that in all these fields due to a single magnet the lines emerge not as if emanating from one point in the iron or steel but as if they came out all over the ends of the bar, as seen in Fig. 1. Thus the thing termed a pole is not a definitely located point of magnetism, but a general spread of magnetiza-

tion all over the iron or steel. Thus the pole is no definitely situated point. This fact is of great importance when it comes to the quantitative study of magnetic forces due to poles. A similar situation arises when the attraction of the earth's gravitational field on a body is considered. Each point of the body is attracted in proportion to its mass. This complex attractive action can be replaced in its behavior by a single force applied at one point, the *center of gravity*. In a similar fashion the diffusely distributed magnetism in the end of a magnet might conceivably be considered to act as if concentrated at a single point in its action with other magnets. Thus we can *idealize* a magnet as having two point magnetic poles located at a distance of l units from each other. The length of the magnet would then be spoken of as l units. Now for some simple geometrical forms it is easy to locate the *center of gravity* by calculation. This is however practically impossible in the case of magnets. Thus the exact location of magnet poles and the length of a magnet is a matter of great uncertainty. This fact complicates the quantitative study of magnets and were it not for a fortunate circumstance discussed on pages 75 and 76 the quantitative treatment of magnetism would be exceedingly onerous.

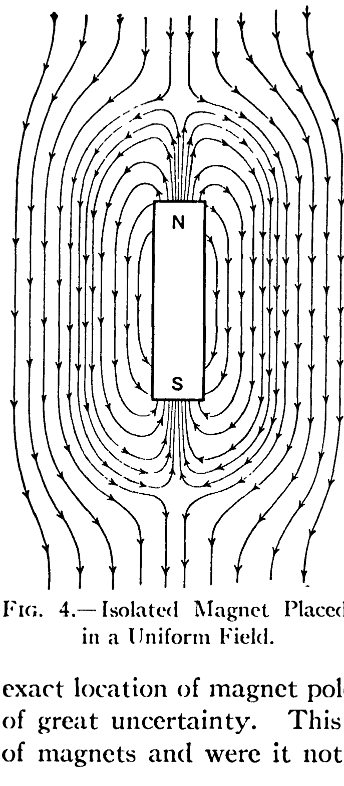


FIG. 4.—Isolated Magnet Placed in a Uniform Field.

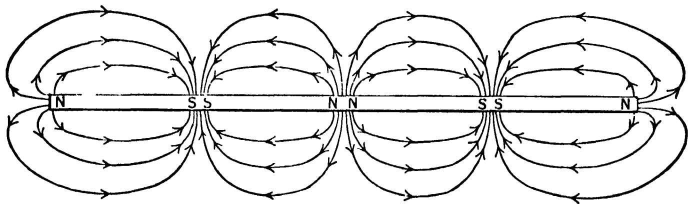


FIG. 5.—Single Magnetized Bar Showing Consequent Poles.

Perhaps one more interesting case may be investigated by the use of iron filings. A steel knitting needle is found to be magnetized on approaching one end to a compass needle in a specific instance with north magnetic polarity. If the other end of this specific needle be

tested it will also be found to exhibit north magnetic polarity. According to the general law of magnetization, however, there should be just one north pole and one south pole in magnet. If the needle be placed on the projection screen and dusted with iron filings as for other magnets the apparent paradox is at once solved. For it appears that there are several poles in the needle as indicated by the iron filings. In fact the pattern of filings appears, as shown in Fig. 5, as if there were in the one needle four separate magnets so arranged that the south poles are opposite the south poles and the north poles are opposite the north poles. It accordingly happens in this one case that with four magnets thus arranged there is a north pole at each end of the rod. The needle does not therefore violate any principles but represents what are termed *consequent* poles. Such a condition is created in a simple fashion by winding a coil about the wire with the sense of winding in the first quarter in one direction (say counterclockwise as viewed from the end), in the next quarter the sense of the winding is in the opposite sense (i.e., clockwise when viewed from the end), at the third quarter the winding is reversed again, and the same

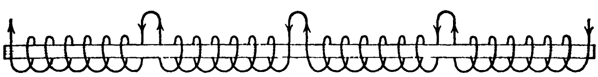
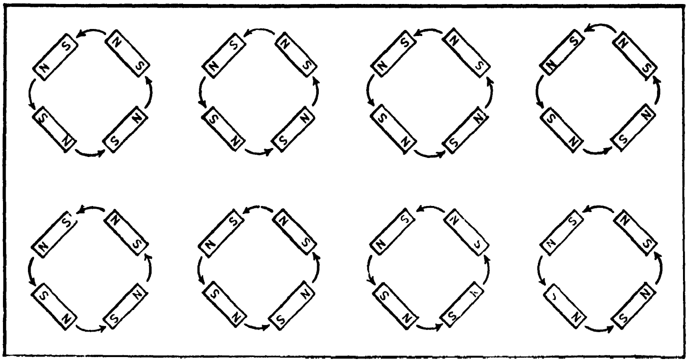


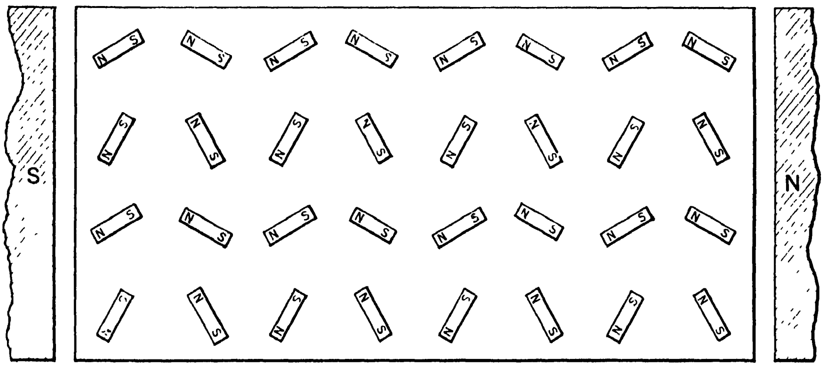
FIG. 6.—Windings for Producing Consequent Poles.

way for the fourth quarter, as shown in Fig. 6. Thus when a current is passed through the wire the consecutive or opposing poles appear.

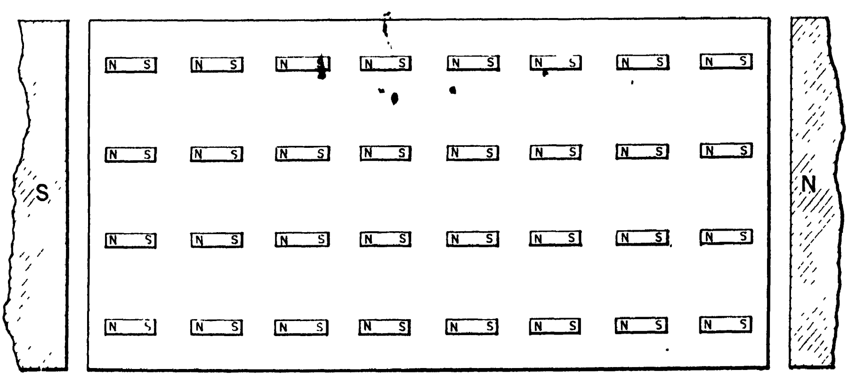
The experiments outlined give one a chance to form a fairly good qualitative picture of magnetism. We can assume that the magnetic substance contains a large number of minute magnetic elements, perhaps the actual molecules, perhaps small crystal units. In the non-magnetized state these units are either disoriented or else oriented in such a way that their poles neutralize each other. Such an arrangement for example is shown in Fig. 7*a*. In a magnetic field this mass of neutralizing magnets showing no sensible external effect is subjected to an external directing force. Depending on the rigidity of their binding these small magnets now tend to orient themselves in the direction of the field, Fig. 7*b*. If the substance is very rigid (i.e., like steel) the resistance is great and large forces are needed to produce an effect, which once produced is relatively stable. In soft iron the magnetic particles swing into line easily and when the field is removed as easily swing back. Thus the nature of retentivity is explained. As the particles swing into line the magnetic poles instead of neutralizing each other now add together and reinforce each other. As the alignment



a. Unmagnetized State.



b. Partial Orientation in a Weak Field.



c. Ideal Saturation in a Strong Field.

FIG. 7.—Idealized Picture of the Process of Magnetization.

becomes more and more perfect the magnet becomes more and more nearly saturated with magnetism, Fig. 7*c*. Induction is thus seen to be the action of the field of one magnet in aligning the magnetic elements in soft iron in the neighborhood. If left to itself a magnet will have its small elementary magnets jostled out of line by the heat motions. The higher the temperature the more violent this disorienting tendency. The same effect is produced by jarring, which in the end produces a heating. Heating a piece of iron in a magnetic field and letting it cool jostles up the neutralizing groups of magnets and permits these elementary magnets to realign themselves with the external field as the metal cools. If a magnet be surrounded with a continuous path of soft iron it will not lose its magnetism due to heat as readily as a bar or other magnet whose poles are free. The reason

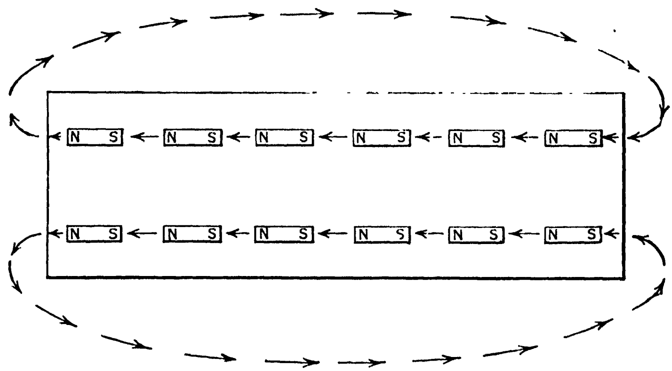


FIG. 8.—Demagnetizing Action of a Magnet's External Field.

is simply that a free magnet has its own field outside in the air opposed to its own internal magnetic alignment. Hence its own external field acts to realign any little elementary magnet that is kicked loose by heat motions in an opposite sense to that which it had in the magnetized state. This is a demagnetizing action of a magnet due to an external field. If the continuous path of iron is present the field in the direction of alignment of the elementary magnets is continuous and in the same sense. Such an alignment also exerts a torque which reacts against the disorientation of any magnetic element in the line. Hence unless the heat vibration is very violent a magnet with a keeper will maintain its magnetism indefinitely. These actions are illustrated in Figs. 8 and 9.

The magnetostrictive phenomena are also easily explained by the qualitative picture given. As the neutralizing groups in unmagnetized iron break up and the magnets tend to align themselves as in Figs.

7*a*, *b*, and *c*, the magnetized bar will lengthen, for the "end on" arrangement of Fig. 7*c* takes more room than the closed arrangement of Fig. 7*a*. As saturation is approached the complete alignment of these magnetic elements brings into play the powerful attractive forces of the magnets on each other in a cumulative manner. Thus lengthwise along the magnetic axis the magnets exert a powerful magnetic attractive force which sensibly shortens the bar.

The nature of the orientation, which is at best never quite complete, and far from it in the case of the permanent magnets, shows why the idealized state of Fig. 7*c* is never reached and why it is that the magnetism appears *distributed* near the ends of the bar instead of appearing at the ends only as in an ideal case.

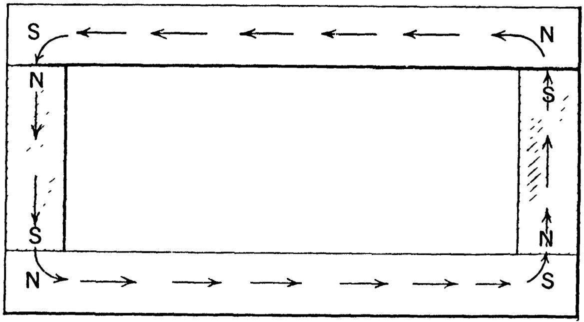


FIG. 9.—Effect of Keepers on Magnets.

This rather complete qualitative picture of the magnetic processes can be roughly verified by filling a glass tube with iron filings and sealing it. If now the tube be placed in a solenoid carrying an electrical current, tapped, and then removed it will be observed to be highly magnetized. As soon as it is shaken the tube returns to its previous unmagnetized state, the shaking breaking the oriented alignment of the filings. A more accurate test of this question was made quantitatively by Ewing and Gans, who pivoted a large number of small magnets in a plane and by applying an external field could observe quantitatively the various phases of a hysteresis loop, described in Chapter XIX, which more accurately characterizes the behavior of a magnet than the qualitative pictures.

CHAPTER III

THE QUANTITATIVE TREATMENT OF MAGNETIC FORCES

6. THE CONCEPT OF THE MAGNETIC FIELD

It was observed from the interactions of magnets that forces must exist between the magnets. Therefore, in the region around a magnetic pole there exists a condition such that another magnetic pole, brought into this region, experiences a force at each point in the region. The direction of the force at any point may be determined by the orientation of a minute compass needle placed at that point, that is, the compass needle so orients itself that it is parallel to the resultant force. It is therefore possible to map the field of force about a magnet by finding the direction which the compass needle would take at all points of the space considered. The field can also be shown by shaking light iron filings over a paper giving the plane in which a study of the field of force is sought. The iron filings in this case constitute minute magnets which align themselves in the direction of the field. *A more rigorous definition of the direction of a field of force would be given by taking the path described by a hypothetical isolated north pole in the field.* It is found that the force at each point, in addition to having a direction, has a definite magnitude, and like all force is therefore a vector quantity. This force varies with the distance from the various magnetic elements which cause the field.

7. DEFINITION OF UNIT POLE, AND POLE STRENGTH, COULOMB'S LAW

In order to determine the magnitude of the field, one must have a unit of measure. The existence of fields of magnetic force is manifested by the forces exerted on magnetic poles. It should therefore be possible to express the magnitude of a pole strength or of a field strength by the force exerted on some other pole. To define this force, one must choose a standard. The choice of a standard can come only from an experimental study of the law of force between poles.

Coulomb was the first to study the forces experimentally. He

initially studied the forces of attraction between poles using the torsion balance. In this study he suspended a magnet from a fiber to which was attached a mirror so that the deflection of the mirror enabled the deflection of the magnet produced by a given force to be measured. The force necessary to produce a given deflection on the balance was known, and hence by measuring the angle of deflection, the force produced by another pole on the suspended magnet could be measured. In his early experiments, he used long magnets so as to study the effect of poles as nearly isolated from their accompanying opposite poles as possible. By this means he found that the force was proportional to what he defined as *the magnetic strength of the pole,* and inversely proportional to the *square of the distance* between the poles. If the one pole be designated as having a strength m and the one brought r cm from it as having a strength m', the force between them can be written $f = \frac{mm'}{r^2}\frac{1}{\mu}$. Here $\frac{1}{\mu}$ is a constant of proportionality and depends on the medium and the units chosen.

The establishment of this law leads at once to a definition of the unit of pole strength. In any system of units, we can *arbitrarily define a unit pole as one which repels an exactly equal pole at unit distance with unit force when* μ is unity. For if $f = 1$, $r = 1$, and $\mu = 1$, then $mm' = m^2 = 1$, and $m = 1$. In physics the C.G.S. system of units is generally used. In this system the unit of force is the dyne, and the unit of distance the centimeter. Hence the *unit pole in the C.G.S. system is the pole which repels an exactly equal pole at the distance of one cm with the force of the one dyne* (*in vacuum*).

Practically it is difficult to obtain exactly equal poles. It is furthermore difficult to locate accurately the *center of magnetism* in a pole so as to measure distances accurately. Thus the measurement of pole strengths must be left to indirect methods. We shall see later that this is accomplished by performing two experiments which give us a field strength and a pole strength multiplied by the distance between the poles from the equations involved. (See Chapter IV.) In static electricity one has an easier time in measuring the electrical quantity q, whose defining equation $f = \frac{qq'}{r^2}\frac{1}{D}$ is very similar to the magnetic equation, as it is easy to realize equal quantities.

As regards the quantity μ, this is a constant involving the nature of the materials in which the two poles are immersed. It is assumed to have the value unity for empty space. For air it is so nearly 1 (1.0000004) that it can be called unity. For iron it is very high (about 10,000). μ is called the magnetic permeability. (See Chapter XVIII.)

The magnetic field in a region can be studied by means of the forces exerted on a test pole. As the force is dependent on the pole strength of the pole used for studying the field, it is convenient to express the field in terms of the force on a standard pole. To this end *we define field strength at a point as the force exerted on a unit pole at that point.* That is, the field strength H is given by $H = \frac{f}{m}$, where f is the force in dynes on a pole of strength m. If $f = 1$ dyne, and m is a unit pole, H is a unit field strength. *Unit magnetic field in the C.G.S. system is therefore the field that exerts 1 dyne of force on a unit pole placed in it.* This unit of field strength has been named the *gauss*. Thus a field of 1 gauss exerts 1 dyne of force on unit pole, and m dynes of force on a pole of strength m. The value of the concept of field strength lies in the fact that if H, the field strength, is known, the force on any pole m is at once given by the value of the product Hm.

8. DIMENSIONS OF MAGNETIC POLE STRENGTH

Before proceeding further, one may digress a minute to determine the dimensions of the new quantity magnetic pole strength. In the equation above, let the two poles be equal. That is, $m = m'$. The equation then becomes

$$f = \frac{m^2}{r^2}$$

whence

$$m = \sqrt{r^2 f},$$

if one disregard the dimensions of μ as will be done throughout the text.

Force has the dimensions given by: $f = M\,L\,T^{-2}$

Also

$$r^2 = L^2.$$

Therefore

$$m = \frac{M^{1/2}\,L^{3/2}}{T}.$$

Since field strength H is the force per unit pole,

$$H = \frac{f}{m},$$

whence

$$H = \frac{M^{1/2}}{L^{1/2}T}.$$

9. UNITS AND DIMENSIONS

The physical units and their applications are organized on the basis of the mechanical system of Newton. All gross mechanical occurrences in nature are related by mathematical relations of a more or less simple nature to the system of definitions or axioms contained in Newton's three laws of motion. The quantitative expression of these laws in a restricted form is the expression defining force, $f = ma$, where m is the mass or inertia of the body and a its acceleration. Acceleration is rate of change of velocity, $\frac{dv}{dt}$, and velocity is defined as the space or distance covered in unit time, or $\frac{ds}{dt}$; so that $a = \frac{d^2s}{dt^2}$, and $f = m\frac{d^2s}{dt^2}$. Thus the fundamental measurable quantities entering into this basic equation descriptive of all physical occurrences are length or distance, time, and inertia or mass. All change in the condition of a mechanical body is accordingly expressible by appropriate combinations of these quantities. Now the units employed for measurement in terms of these three fundamental quantities are defined in terms of simple convenient arbitrary standards, the second, the centimeter and the gram. The second is merely a convenient fraction of the mean solar day; it is $\frac{1}{60} \times \frac{1}{60} \times \frac{1}{24}$ of the mean time that elapses between successive transits of the sun across the meridian. Why this fraction is used instead of a fraction expressed in the more convenient decimal system is hard to say. The antiquity of this time system, going back to the ancients, and the fact that our custom has been to use it in tables for so many years, are the probable causes of its continuation. The unit of length is also a purely arbitrary, convenient standard. When the standard meter bar was made, it was supposed to be one forty millionth the circumference of the earth through the poles. However, later measurements showed this value to be incorrect and the standard meter is now merely the length of a platinum bar kept at the International Bureau of Weights and Measures at Sèvres, France. Michelson later determined the length of the standard meter in terms of the wave length of three lines in the spectrum of cadmium so that we possess a check on the standard should it vary. Finally, the unit of mass was for convenience chosen in terms of the inertia of a cubic centimeter of water at 4° C, the temperature of its greatest density.

As is seen, the choice of our so-called fundamental units indicates

that they are in *no sense fundamental*, but are merely convenient, or even chance arbitrary standards to use with our Newtonian system of units. It is not surprising then to find that our real and probably truly fundamental units, such as the electron, h, the Planck action quantum, the mass of the electron and of the hydrogen nucleus, are expressed as peculiar odd ratios of our chosen units in the fields where they are measured. We are accordingly committed to expressing all phenomena in physics in terms of Newtonian mechanics and ultimately in terms of these three arbitrarily chosen fundamental units in the C.G.S. System. Leaving, however, the units aside, we in general can express by Newtonian mechanics all phenomena in terms of powers of length, mass and time.

To better understand the foregoing statements, we may discuss the process of obtaining a new entity or quantity in nature. When a new phenomenon of nature is observed the practice is to deduce quantitatively the behavior manifested by means of controlled, quantitative investigation and to formulate behavior in terms of a mathematically expressed law. For example, Coulomb in deriving the law of electrostatic force essentially proceeded as follows. He took two charges and actually studied how the force varied for the same two charges as distance alone was varied. Then keeping distance constant he varied first the state of charge on one body and then the state of charge on the other. The numerical results or data were set down in table form and by analysis the mathematical laws controlling them were deduced and expressed in the generalized form of an equation. This procedure led at once to the idea that the force between two electrified bodies depended inversely on the square of the distance and on the product of two terms that varied with the electrical state of excitation of the two bodies. Either one of the two quantities q and q' describing the two electrical states of the bodies in the law deduced, $f = \frac{qq'}{r^2}$, he defined as a new quantity in nature, a property of electrification, and called it the *quantity of electricity*.

The law governing a phenomenon being once formulated, and the new quantity in nature defined in terms of things measurable in the mechanical terminology based on Newton, the unitary value of this quantity may easily be defined. For assuming as *fundamental for practical purposes* the C. G. S. system of units defined above, all that is needed is to solve the equation discovered for the quantities under simplifying assumptions, and to set the quantity as *unity when each of the items in the equation is taken as unity on the C. G. S. system.* Thus, one would write $qq' = fr^2$, and simplify it by letting $q = q'$;

then $q^2 = fr^2$, and $q = \sqrt{fr^2}$. Unit electrostatic quantity was therefore taken as that quantity for which $\sqrt{fr^2}$ equaled unity as applied to the phenomena investigated. Put into words the law then is the formal definition so often learned in a parrot-like fashion by most students without an understanding of its meaning. This process of becoming familiar with a new concept of physics by means of the defining equation, and defining the unit in the manner outlined above should give a far clearer idea of the concept and the unit, as it presents in concise mathematical form the relations involved. This method will constitute the procedure to be followed throughout the book in defining the many new units to be encountered.

Now in developing a science numerous such new quantities are found and it becomes essential to relate and correlate them with each other in order that they may be used to the best advantage. It is further useful, so to speak, to "keep books" when new quantities are found to make sure that things equated are really legitimately equated. Thus writing an expression equivalent to an energy equal to something that is not energy would introduce obvious errors into our results. Finally some new quantities are derived under conditions where their nature is not obvious and it pays to establish their nature to make the best use of our newly gained knowledge. To avoid possible mistakes, and to work to the best advantage, we can check our equations by analyzing them into the component three fundamental elements underlying all Newtonian mechanics; that is we determine the *dimensions* of a quantity in terms of length written L, mass written M and time written T. If the quantities on the two sides of an equation have outside of numbers or numerical ratios, which we ignore, the same powers of L, M and T, the equations are dimensionally correct. By the same process we can determine the dimensions of a new quantity in terms of known ones.

The dimensions of a few important physical quantities are as follows: L = length, M = mass, T = time; velocity, $v = \frac{s}{t} = LT^{-1}$, acceleration, $a = \frac{v}{t} = LT^{-2}$, force, $f = ma = MLT^{-2}$; work, $w = fs = ML^2T^{-2}$, power, $p = \frac{w}{t} = ML^2T^{-3}$. Now two equations frequently written are that impulse, ft = momentum, mv, and that kinetic energy, $\frac{1}{2}mv^2$, = work, fs. To test the correctness of these assertions we can set $ft = (MLT^{-2})T = M(LT^{-1}) = mv$ which we see is an identity, and $\frac{1}{2}mv^2 = \frac{1}{2}M(LT^{-1})^2 = (MLT^{-2})L = fs$, which, neglect-

ing a numerical constant, is also seen to be an identity. Another example of the use of dimensions is from modern physics. It was found that a doby can be set into vibration at its natural frequency ν only if it receives an energy $\frac{1}{2}mv^2$ given by $h\nu$, where h is a new universal constant. One may ask what units h is to be expressed in, or in other words what h is dimensionally. $\frac{1}{2}mv^2 = ML^2T^{-2} = h\nu$, now ν = frequency = number per unit time = T^{-1}. Thus $ML^2T^{-2} = hT^{-1}$, or $h = ML^2T^{-1}$. Now ML^2T^{-2} is energy so that h has the dimensions of energy times time and h is evaluated as 6.55×10^{-27} *ergs* $\times$ seconds.* Again we can consider temperature as an admirable illustration. Temperature begins to acquire a physical significance with the adoption of the ideal gas law relation $pV = RT$, where T is absolute temperature. Now $p = \frac{f}{A} = ML^{-1}T^{-2}$, $V = L^3$, and $pV = ML^2T^{-2}$, or energy. Hence, RT is work or energy, and in fact the later development of the kinetic theory showed that RT is $\frac{2}{3}$ the total kinetic energy of the molecules of a gas. Thus we gain an actual understanding of the nature of RT and hence of temperature, for T multiplied by R gives the energy in the gas at a temperature T, T being the scale factor of the energy content.

Another very interesting, important and illuminating illustration comes from electricity.† In the electrostatic system quantity is defined as $q = \sqrt{fr^2D}$, where D is a new constant of the materials in which q finds itself, called the dielectric constant.‡ Its dimensions are unknown and are ignored in the text except in this discussion. Thus $q_{ES} = D^{\frac{1}{2}}M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}$. Now in the electro magnetic system of units $q_{EM} = it$, where i is current and t is time. But i is defined from Ampère's law which says that $f = \frac{idsm}{r^2}$, where ds and r are lengths and m is magnetic pole strength. Now $f = \frac{mm'}{\mu r^2}$ where m is pole strength, r a distance and μ is the magnetic permeability, a property of the magnetic materials surrounding m, the dimensions of which are unknown. The dimensions of μ are again ignored in the text

* Again moment of momentum $mvr = ML^2T^{-1} = h$, so that h has simultaneously the properties of mvr and wt, a conclusion of far-reaching importance for atomic structure.

† This discussion is included at this place for reference. It is appropriate here as an example of one of the uses of dimensional reasoning and should somewhere be included in the text in regard to the ratio of units. It was deemed best to insert it at this point as there seems no other suitable place in the text.

‡ In the use of Coulomb's law on page 67, D was omitted as it was not discovered until much after Coulomb's time by Faraday.

except in this discussion. Accordingly, $m = \mu^{\frac{1}{2}}M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}$. Thus $i = \dfrac{ML^2T^{-2}}{\mu^{\frac{1}{2}}M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}} = \mu^{-\frac{1}{2}}M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}$, and $it = \mu^{-\frac{1}{2}}M^{\frac{1}{2}}L^{\frac{1}{2}}$. Now the quantities q_{EM} and q_{ES} should be dimensionally the same, for they represent *quantity of electricity,* differently measured and yet believed to be the same. Thus $q_{EM} = q_{ES}$ and $D^{\frac{1}{2}}M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1} = \mu^{-\frac{1}{2}}M^{\frac{1}{2}}L^{\frac{1}{2}}$. This relation can only hold true if $LT^{-1} = \dfrac{1}{\sqrt{\mu D}}$. Now LT^{-1} is a velocity v. Thus $\dfrac{1}{\sqrt{\mu D}}$ must be a velocity. The question then arises as to what this velocity represents. If the unit q_{EM} be measured and the number of the units of q_{ES} equivalent to one unit q_{EM} be determined experimentally the experiments establish the relation as 3×10^{10} electrostatic units of quantity, q_{ES}, being equal to 1 electromagnetic unit, q_{EM}. The numerical ratio is the velocity of light in empty space in cm/sec, i.e., absolute C.G.S. units. Hence $\dfrac{1}{\sqrt{\mu D}}$ might be expected to represent the velocity of an electromagnetic, or light wave. It was the discovery of this ratio that led Maxwell to investigate the electromagnetic relations and deduce the fact that they should lead to a wave motion in empty space with the properties of light and a velocity $\dfrac{1}{\sqrt{\mu D}}$. That that quantity $\dfrac{1}{\sqrt{\mu D}}$ actually does represent the velocity of light can be shown as follows from our dimensional analysis.

If we let N_{ES} be the *number of E.S.U.* in a given quantity of electricity the expression $N_{ES}(M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}D^{\frac{1}{2}})$ is the complete expression for the quantity, and if N_{EM} be the number of E.M.U. in the same quantity, $N_{EM}(M^{\frac{1}{2}}L^{\frac{1}{2}}\mu^{-\frac{1}{2}})$ is also the complete expression for the quantity. Here N_{ES} and N_{EM} are mere *numbers giving the numerical values involved.* Thus since the two expressions for quantity above represent the same quantity,

$$N_{ES}(D^{\frac{1}{2}}M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}) = N_{EM}(\mu^{-\frac{1}{2}}M^{\frac{1}{2}}L^{\frac{1}{2}})$$

$$\frac{N_{EM}}{N_{ES}} = \frac{1}{3 \times 10^{10}} = LT^{-1}/\frac{1}{\sqrt{\mu D}}.$$

As LT^{-1} in units is in absolute C.G.S. units $= 1$ cm/sec, and $\dfrac{1}{\sqrt{\mu D}}$ is related to this in the ratio of 3×10^{10} to 1, for N_{EM} and N_{ES} are pure numbers, hence $\dfrac{1}{\sqrt{\mu D}}$ must be a velocity 3×10^{10} times as great as

T^{-1}. Thus $\dfrac{1}{\sqrt{\mu D}}$ represents a velocity of 3×10^{10} cm per second r that of light. Hence by dimensional reasoning, coupled with xperimental fact, we have arrived at the conclusion that while sepa-ately μ and D have indeterminate dimensions, in the form $\dfrac{1}{\sqrt{\mu D}}$ they epresent a velocity which is that of propagation of electromagnetic aves through empty space or 3×10^{10} cm/sec, and that the ratios f the *fundamental* electrical units in use in this course based on this eduction are equal to the velocity of light in magnitude.

10. APPLICATION OF LAW OF FORCE TO SPECIAL CASES

Now Coulomb's early measurements were inexact because of the omplication of other poles, and the exact study of the forces between oles and their measurements must be achieved in a more indirect nanner. Since we determine pole strength by forces we can use the nethods applicable to forces to study poles. The magnetic force as magnitude and direction. It is, therefore, a vector. The study f the resultant force at any point, due to a series of magnetic poles n its neighborhood, is accomplished through the vectorial addition f the separate forces due to the separate poles. This treatment *ssumes that the force due to each pole acts independently of the presence f other poles.* Using this assumption, we can study the field due to ertain arrangements of poles which have a practical bearing on roblems to come.

Case 1.—The field of an isolated north or south pole of strength $\pm\, m$ in air at a point r cm distant from the pole is by definition

$$H = \pm \frac{m}{r^2} \text{ in the direction of } r.$$

Therefore the force on a pole m' is $m'H = \pm \dfrac{mm'}{r^2}$.

Case 2.—The field at a point A distant d cm from the center C, Fig. 10, of a bar magnet whose ength is l cm, and whose pole trength is m units, may be found s follows. Due to the north pole he force is one of repulsion at A nd is expressed by

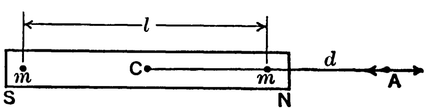


FIG. 10.

$$f_1 = + \frac{m}{\left(d - \dfrac{l}{2}\right)^2}, \text{ on unit pole,}$$

and the force due to the south pole is one of attraction and is expressed by:

$$f_2 = -\frac{m}{\left(d + \frac{l}{2}\right)^2}, \text{ on unit pole.}$$

The resultant force f_r is given by:

$$f_r = +\frac{m}{\left(d - \frac{l}{2}\right)^2} - \frac{m}{\left(d + \frac{l}{2}\right)^2}, \text{ on unit pole.}$$

Reducing this by algebraic manipulation one arrives at the final expression

$$f_r = \frac{2ldm}{\left(d^2 - \frac{l^2}{4}\right)^2}, \text{ on unit pole at } A.$$

If l is small, l^2 may be neglected compared to d^2. The force on unit pole at A then becomes

$$f_r = \frac{2ml}{d^3} = H$$

which suffices for the solution of many simple problems.

It is seen from this that the field falls off rapidly with the distance from a magnet, and it makes the detection of the magnet at a large distance very difficult. Thus, although it was suggested that the magnetization due to the electrical equipment of a submarine might be used for detecting its presence, the very weakness of the magnetic forces at distances of approach which made the submarine dangerous, made detection impossible.

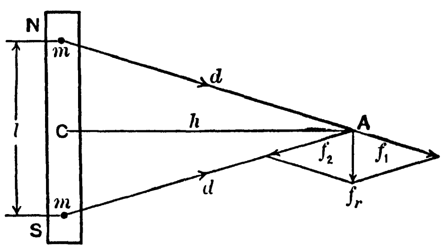


FIG. 11.

Case 3.—Consider the force at a point A, Fig. 11, distant h from the center C of a bar magnet of a length l, where h is taken in a direction perpendicular to the bar magnet of pole strength m at its center. Call d the distance from the poles of the magnet to the point at which the force is to be determined. It will be noticed that the distances d are equal. In this case the force

exerted on unit north pole by the north pole will be away from the latter and will have a magnitude f_1 given by

$$f_1 = \frac{m}{d^2} = \frac{m}{h^2 + \frac{l^2}{4}}.$$

The force on unit north pole due to the south pole will be directed toward that pole and will have the magnitude f_2 given by

$$f_2 = \frac{m}{d^2} = \frac{m}{h^2 + \frac{l^2}{4}}.$$

In this case the resultant force f_r is obtained by the vectorial addition of f_1 and f_2. It is represented by the small vector f_r, and since f_1 and f_2 are equal in magnitude, f_r will be parallel to the axis of the magnet. From similar triangles it can be seen that

$$\frac{f_1}{f_r} = \frac{d}{l}, \quad f_r = f_1 \frac{l}{d} = f_2 \frac{l}{d}$$

whence

$$f_r = \left(\frac{m}{h^2 + \frac{l^2}{4}}\right)\left(\frac{l}{\sqrt{h^2 + \frac{l^2}{4}}}\right) = \frac{ml}{\left(h^2 + \frac{l^2}{4}\right)^{\frac{3}{2}}}, \text{ on unit pole at } A.$$

If l is small compared to h

$$f_r = \frac{ml}{h^3}, \text{ on unit pole at } A.$$

Case 4.—For any other point A around the bar magnet the resultant force is easily determined by the following method. From A, Fig. 12, draw a line AD parallel to the axis of the magnet, join A to the north and south poles, call the force of repulsion due to the north pole f_1, and that of attraction due to the south pole f_2. Call l the length of the magnet, and draw a perpendicular D to the center C of the magnet intersecting AD at D.

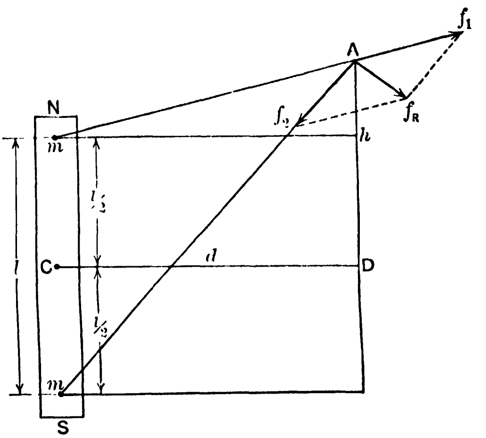


FIG. 12.—Calculation of Field at Any Point Near an Isolated Bar Magnet.

If one call d the distance CD from the center of the magnet along the normal to the point where it intersects the line from A parallel to the axis, and if one call h the distance AD from this point of intersection to the point A, one can express the forces f_1 and f_2 on unit pole at A in terms of the pole strength m and the distances h and d by the equations:

$$f_1 = \frac{m}{d^2 + \left(h - \frac{l}{2}\right)^2}$$

$$f_2 = \frac{m}{d^2 + \left(h + \frac{l}{2}\right)^2}.$$

The resultant force f_r on unit pole at A may be at once found from the relation

$$f_r^2 = f_1^2 + f_2^2 + 2f_1f_2 \cos \theta$$

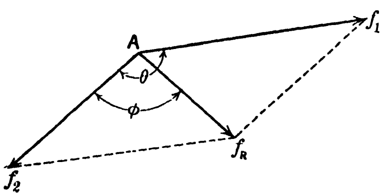


FIG. 13.

where θ is the angle between f_1 and f_2. It will be noted that if θ is less than 90° the cosine term is positive, and if θ is greater than 90° it is negative. The angle φ of the resultant f_r with one of the original forces f_2, Fig. 13, can be found, as in the solution of all force triangles, from the well-known relation that

$$\frac{f_1}{f_r} = \frac{\sin \varphi}{\sin \theta}, \text{ whence } \sin \varphi = \frac{f_1}{f_r} \sin \theta$$

and the value of φ can be found from the tables.

Case 5.—For any number of magnets the problem resolves itself into calculating the forces for each pole and composing the forces to a single resultant, using the laws of composition of forces.

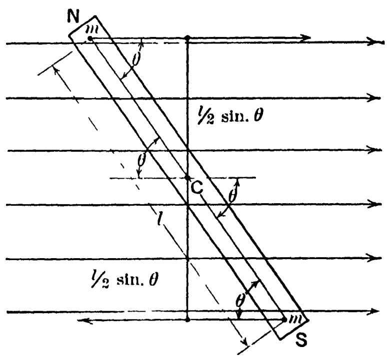


FIG. 14.—Forces on a Magnet in a Uniform Field.

Case 6.—The next problem which is of importance is the interaction between a field and a magnet oriented in any position in that field. Assume a uniform field of strength H represented by the parallel lines of Fig. 14. Assume a magnet of pole strength m lying in the field in such a way that its axis makes an angle θ

with the field. The north pole is urged in the direction of the arrow at N. The south pole is urged in the direction of the arrow at S. The result of the action of these two forces, since they must be equal, the field being assumed uniform and the poles equal, will be to exert a simple torque on the magnet, for the equality of the forces in opposite directions on the magnet as a whole precludes translatory motion. This torque would cause the magnet to turn so that its north pole is towards the south pole creating the field, and its south pole is towards the north pole creating the field. It is important to be able to calculate the magnitude of this torque. The force on the north pole is by definition Hm. If l is the length of the magnet, it acts on a lever arm $\frac{l}{2} \sin \theta$. The south pole is urged in the opposite direction with a force Hm and the force moment which would cause rotation in the same sense as that acting by means of the north pole will be $Hm \frac{l}{2} \sin \theta$. The resulting force moment symbolically expressed by the quantity G is given by

$$G = 2\, Hm \frac{l}{2} \sin \theta = Hml \sin \theta.$$

It is seen that the torque depends on the field strength, on the angle θ, and on the product of the pole strength by the length of the magnet. This quantity *ml is a constant of the magnet as long as its pole strength is unchanged. We define this product ml as the magnetic moment of the magnet and denote it by the symbol M.* The use of the moment M of a magnet is exceedingly convenient inasmuch as it enables us to use a quantity characteristic of the properties of a given magnet which is independent of the uncertain length of the magnet as is seen in Chapter I. The magnetism at the end of a magnetic bar of iron is distributed in a very complex manner. The determination of the location of the center of magnetism and hence determination of the distance between the poles composing the magnet is practically impossible with such a distribution. The magnetic moment defines product of the pole strength times the length of the magnet, which is independent of any knowledge of the distribution of magnetism, but characterizes the action of the magnet on other magnets and in fields. Thus all our equations involving the study of magnets involve simply the evaluation of this convenient constant which sufficiently defines a magnet for *practical* purposes. The torque on the magnet in the field is accordingly written as $G = HM \sin \theta$. This equation is very important and will be used freely in what follows.

CHAPTER IV

THE ABSOLUTE DETERMINATION OF POLE STRENGTH AND MAGNETIC FIELDS. THE EARTH'S FIELD—FIELD CONVENTIONS

11. THE ABSOLUTE DETERMINATION OF POLE STRENGTH AND MAGNETIC FIELDS

WE now turn to an important question: the method of measuring pole strength. As was stated at the beginning, it is impossible to make direct use of the definition of pole strength in order to measure that quantity. It is essential, however, for us to be able to determine the pole strength m, or the magnetic moment M, in absolute units. If we could determine either the pole strength of a magnet or the magnitude of a uniform field H in absolute units it would be possible thereafter to determine the strength of all other magnets by means of these, using the equation for the torque on a magnet in a uniform field, deduced in the last Chapter, or the tangent law to be deduced in Chapter V.

As it originally was impossible to obtain a magnetic field of known value or a magnet of known pole strength it became necessary to measure these quantities indirectly. The method to be outlined makes use of the fact that the earth in the absence of magnetic materials gives a practically uniform magnetic field. By studying the forces acting on a magnet in such a uniform magnetic field it is possible to determine both the value of the magnetic moment M of the magnet and the absolute value of the earth's field H.

Because there are two unknowns, M and H, to be determined, this requires that two simultaneous equations containing M and H be set up, solution of which will give M and H. This in turn requires two experiments which give two relationships between M and H. The first experiment gives the ratio of the magnetic moment M of the magnet to the strength of the earth's field H, or the quantity $\frac{M}{H}$. The second experiment gives the product of the magnetic field H and the magnetic moment M of the magnet. From the value of these

two ratios M and H may be solved for, as we have two simultaneous equations with only two unknowns.

12. EXPERIMENT I, M/H

The magnet of pole strength m, length l, and moment M, Fig. 15, is placed with its axis at right angles to the earth's magnetic field H. At a distance r from the center of the magnet there is placed a small compass needle represented in the figure by magnet A. As a result of the field due to the bar magnet, the small compass needle suffers a torque tending to pull its south pole toward the magnet and to repel

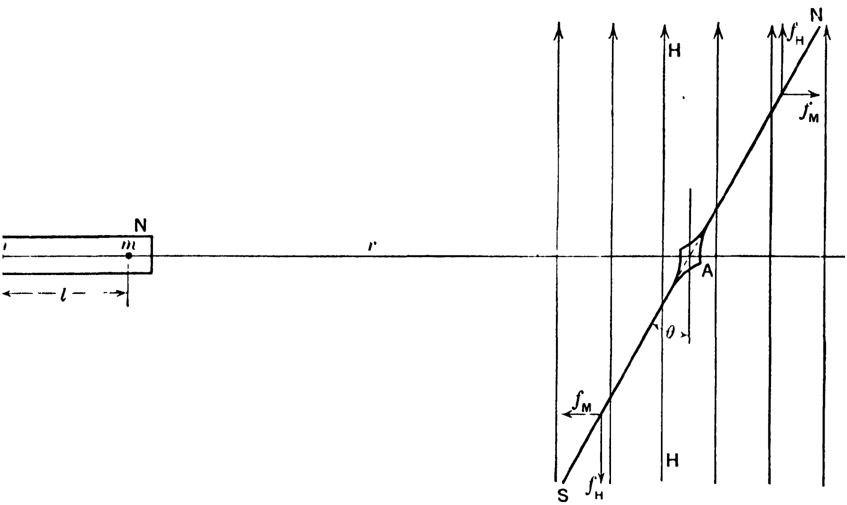


FIG. 15.—The Magnetometer, Experiment I, M/H.

the north pole. At the same time, in the earth's field the small compass needle suffers a pull on its south pole urging it downward in the figure, and on its north pole urging it upward in the figure. The force f_M acting on the north pole of this needle due to the bar magnet and the force f_H acting on the same pole tend to cause a rotation of the compass needle in opposite senses. Since the torques vary with the angle θ, the needle of the small compass will come to rest at an angle θ with the field such that the torques due to the earth's field and to the bar magnet are equal.

The conditions for equilibrium may then be determined. G_M represents the force moment on the compass needle pulling its south pole toward the bar magnet. Assuming the magnetic moment of the small compass needle to be M', and its pole strength to be m', the

torque due to the field of the bar magnet at A which we will designate by $G_M = \frac{1}{2} f_M M' \cos \theta$. In a similar fashion the torque produced by the earth's field H due to its action on the south pole of the small compass needle is given by $G_H = \frac{1}{2} f_H M' \sin \theta$. Since the north pole is acted on in the same fashion the torques on the small compass needle of moment M' will be twice the values given above. In equilibrium $G_H = G_M$. Therefore $f_M M' \cos \theta = f_H M' \sin \theta$. But $f_H = m'H$, therefore,

$$\frac{f_M}{m'H} = \tan \theta.$$

Since f_M, the force on a pole of strength m' at a point A due to a magnet of moment M, is approximately given by $\frac{2Mm'}{r^3}$ as seen in the last chapter, therefore,

$$\frac{M}{H} = \frac{1}{2} r^3 \tan \theta.$$

13. EXPERIMENT II, MH *

In order to get a value of MH the bar magnet used above is suspended in a stirrup in the earth's field by a fine fiber so that the bar magnet rotates about a vertical axis through its center, and perpendicular to its length. The fiber must be so fine that its torsional constant is negligible compared to the torques acting. If such a magnet is twisted from its position of rest parallel to the earth's field a torque is set up tending to urge it to return to its position of rest. By the last chapter this torque is given by

$$G = HM \sin \theta.$$

For small values of θ (less than $10°$), $\sin \theta$ approximates θ, whence one has,

$$MH = \frac{G}{\theta}.$$

* An alternative way of determining the product MH is to suspend the bar magnet of moment M perpendicular to the earth's field by a bifilar suspension and observing the angle through which the magnet is deflected from the rest position by means of a mirror and scale. The determination of the rest position of the magnet and the orientation of the axis of the magnet perpendicular to the earth's field with a bifilar suspension is achieved by replacing the magnet by a dummy magnet of the same dimensions, made of a non-magnetic substance. The product MH is then evaluated from the angle of deflection and the characteristic equation for bifilar orientation. This method is, however, much less accurate than the method of oscillations which is the method in standard use.

Now if $\frac{G}{\theta}$ is known we can get the product MH. $\frac{G}{\theta}$ is nothing else than the torque constant of the system for oscillatory motion. If the fiber suspending it exerts no influence on the period of oscillation, a well known theorem in mechanics relates this force constant for oscillatory motion to the period of oscillation T of the magnet and its moment of inertia I.
This is

$$T = 2\pi\sqrt{\frac{I}{\frac{G}{\theta}}}$$

Therefore

$$T = 2\pi\sqrt{\frac{I}{MH}}$$

and

$$MH = \frac{4\pi^2 I}{T^2}.$$

By experiment I, however, we had

$$\frac{M}{H} = \frac{r^3 \tan \theta}{2}.$$

Solving the two equations we obtain

$$H^2 = \frac{8\pi^2 I}{T^2 r^3 \tan \theta}$$

and

$$M^2 = \frac{2\pi^2 I}{T^2} r^3 \tan \theta.$$

Thus we have obtained the value of M the magnetic moment of the bar magnet and the value of H the earth's field by the two measurements. This is the fundamental measurement of the magnetic quantities in terms of the absolute C.G.S. system of units. The method is therefore the basis of all our quantitative knowledge of magnetic phenomena and as such is of prime importance. Having once determined either M or H the evaluation of any other magnetic poles or fields is a simple matter by use of the tangent law, or the torsion balance. (See end of Chapter V.) The instrument used in making the determination of MH is known as the *magnetometer* and it is used quite widely in studying the earth's magnetic field. The magnetic standards having once been established by magnetometer methods we are able to calibrate magnetic fields produced by currents

so that in terms of our current standards we are now able to produce magnetic fields of known value without recourse to the investigations outlined.

14. NOTE ON THE MEANING OF THE TORSIONAL CONSTANT T_0

It is often useful in the study of fields to use a suspension where the torsional constant is not negligible. If this is the case the mechanical torque G on a fiber twisted through an angle θ is given by $G = T_0\theta$. T_0 can be computed from the dimensions of the fiber by the equation $T_0 = \frac{\pi r^4 \eta}{2l}$, where r is the radius of the fiber, l is its length, and η is the coefficient of rigidity to be found in tables for any substance. η is given in absolute units, and if r and l are in *cm*, then T_0 is in dynes $\times$ *cm*, per radian. G is in dynes $\times$ *cm* when θ is in radians.

If T_0 cannot be computed, it can be measured by putting a non-magnetic mass of known moment of inertia I on the suspension and measuring its period of oscillation T. Then as

$$T = 2\pi\sqrt{\frac{I}{\frac{G}{\theta}}},\quad T = 2\pi\sqrt{I/T_0},\ \text{and}\ T_0 = \frac{4\pi^2 I}{T^2}.$$

15. THE EARTH'S MAGNETIC FIELD

The knowledge of the earth's magnetic field which we are able to obtain by the magnetometer is of considerable importance in many domains. The magnetic field of the earth has been measured over a greater part of the earth's surface and is constantly being measured at certain fixed stations. This knowledge of the earth's magnetic field, in addition to helping to answer questions of theoretical interest such as the origin of the earth's magnetic field, has been an aid to navigation. Practically all of the merchant ships still use the magnetic compass, and accurate knowledge of the magnetic field of the earth is therefore essential in the making of marine charts. To some extent the so-called gyrostatic compasses are replacing the magnetic. Their installation is, however, very expensive and a failure of the gyrostatic compass leaves a ship without means for proper navigation. There is no ship in general service that is not equipped with an auxiliary magnetic compass, and the gyro-compasses are checked daily and even hourly against the standard magnetic compasses when under way, owing to the tendency of the mechanically complex gyro-compass to get out of order and undergo a precessional motion which over a course of time causes a progressively greater error in its indication.

The direction of the magnetic field of the earth with respect to the earth's surface varies with the point on the earth's surface chosen. At or near the equator the lines of force run parallel to the earth's surface. This, however, is not the case at the earth's poles. In fact, even in Berkeley, the compass needle if suspended on a horizontal axis would dip downward at an angle of about 70° with the horizontal. Furthermore, the magnetic needle does not point true north, that is, the magnetic axis of the earth is not the same as the axis of rotation. The location of the magnetic north pole is given by

$$N = 70°5' \, N \text{ lat}, \; 96°46' \, W \text{ longitude}$$

and the location of the south magnetic pole is given by

$$S = 72°25' \, S \text{ lat}, \; 155°16' \, E \text{ longitude}$$

Thus it is seen that not only does the magnetic axis of the earth deviate materially from its axis of rotation but the magnetic axis does not even pass through the center of the earth. Furthermore, the angle between the true north and the magnetic needle is not constant from year to year. For instance, it was 17°0′ west of true north at London in 1894. In 1910, it was 16°5′ west of true north at London. In Berkeley, the magnetic north is 18°20′ east of true north.

The deviation of the magnetic needle from the true north is called the *declination*. The inclination of the needle, mounted on a horizontal axis, with the horizontal is in Berkeley about 70° and is called the *dip* or *magnetic inclination*. As stated before, the dip is 0 at the equator and approaches 90° near the poles. The intensity of the field H, as defined and measured above, acting on our compass needles which are mounted horizontally is therefore not the total intensity of the earth's field, for the field acts at an angle θ with the horizontal. The H mentioned in our discussion early in the chapter is the *horizontal component of the earth's field*. If we know the horizontal component H and the angle of dip, θ, the total magnetic intensity of the earth's field H_m may be obtained by the relation

$$H_m = \frac{H}{\cos \theta}.$$

For the use of mariners and students of the earth's magnetic field maps have been constructed to show:

(*a*) Lines of equal declination called *isogonic* lines.
(*b*) Lines of equal dip called *isoclinic* lines.
(*c*) Lines of equal intensity known as *isodynamic* lines.

16. CONVENTIONAL REPRESENTATION OF FIELDS OF FORCE

The representation of the fields of force in terms of such maps is more useful for the mariner than for the physicist. It is often convenient for the physicist to represent a magnetic fielddia grammatically. To do this he has adopted a certain convention. *A unit magnetic field, where the field is one gauss, is depicted by drawing one line per sq. cm of area taken normal to the line representing the direction of the force.* If the field has an intensity of H gauss, there are H lines per sq. cm of area taken normal to the direction of the lines. For instance, if one wished to represent the horizontal intensity of the earth's field by lines of force, which at Berkeley is .25 gauss, one would do this by drawing one line to every 4 sq. cm of area normal to the direction of the lines of force.

Another consideration is also of importance. The field intensity at a distance of 1 cm from a single pole * m is by definition m gauss. Therefore, for the spherical surface of radius 1 cm about this pole there are m lines for each sq. cm of the surface. Now such a sphere has a surface area of 4π sq. cm. Thus, from the pole of strength m there must emerge a total of $4\pi m$ lines of force. *That means that from each unit pole there are 4π lines of force emerging.* The lines of force emerging from a pole are constant in number. Thus 1 cm from a unit pole there is one line of force per sq. cm; two cms from a unit pole there is one line of force per 4 sq. cm. This decrease in the number of lines per unit area as one recedes from the poles indicates the decrease in the field strength as one moves away from the pole.

Sometimes it is more convenient to deal with a tube of force. A tube of force may be considered as the surface surrounding a given number of lines of force issuing from a pole in such a manner that no lines of force emerge through the sides of the tube. It is really the region in space which contains a constant number of lines of force which originally start from a pole. The lines of force near the sides of the tube thus are parallel to the sides of the tube, for they do not emerge. Putting it in another way, one might state that it is a volume parallel to the lines of force which also has in it a constant number of lines of force. The magnetic intensity varies inversely as the area of cross section of such a tube of force. *We may define the flux density in such a tube as the number of lines of force divided by the area taken normal to the lines.*

* Of course, it is understood that such an isolated single pole is an idealized concept, for a pole can never exist isolated by itself. It is always accompanied by an equal and opposite pole in the same magnet.

CHAPTER V

ELECTRIC CURRENTS

17. HISTORICAL SKETCH

THE discovery of electrical currents dates from Gray and du Fay in 1729, who showed that a static electrical charge was carried from one body to another by means of conducting wires. Before 1750 it was known that the velocity of transport of electricity was very high. No further progress could be made in the study of currents owing to the fact that the currents from static charges flowed over such short intervals and were so weak that experimentation was impossible. In the period 1786 to 1799 the discovery of the means of producing larger currents was made. It came as a result of the investigations stimulated by the physiological researches of Galvani. He had observed that a static machine on a table made frog's muscles twitch. Franklin's experiment with the kite had shown the identity of static electricity and lightning. Galvani therefore hung muscles on wires in the air. As might have been expected from Franklin's work, Galvani observed twitches. However, twitches occurred in the absence of thunder showers. Galvani found that a single wire in some cases sufficed to cause the twitching. He explained the effect as coming from the muscle nerve system as such effects had been observed in the electrical fish. Volta was keener. He ascribed the effect to the metal. He found that the effect was strong and reproducible when wires from two metals in contact were touched to two parts of the nerve. He found that one metal sufficed if there were a temperature gradient in the metal. The effect was made still more pronounced if the two metals were separated by a damp cloth containing an electrolyte. This was the origin of the Voltaic pile, or electric cell. Volta showed by means of an electroscope and multiplier (see page 210) that the effect was produced by an electrical charge. He found that if he took a series of disks of two different metals separated alternately by wet cloths he obtained an additive effect so that the electrical effect of one element was multiplied by the number of elements used. Such sources of electrical current

enabled currents to be investigated. Today, we have the following sources of electrical currents:

(*a*) The flow of static electricity.

(*b*) The Voltaic pile.

(*c*) Heating of a junction of two metals.

(*d*) The cutting of magnetic lines of force by a conductor (Faraday, 1831).

(*e*) Animal electrification, which is merely a manifestation of Voltaic pile activities in organic tissues.

18. MAGNETIC FIELD OF A CURRENT

Until 1819, no quantitative measurements of electric currents were possible. Oersted was the discoverer of a phenomenon which lead to a means of measuring the currents. Oersted, having noticed the electrical polarity of the Voltaic pile, looked for a magnetic effect because magnets were known to have polarity. In experimenting with the circuits of electricity he observed that when an electrical circuit from a cell was closed a magnetic needle near one of the conducting wires was deflected. The investigation showed that there was a magnetic field about a conductor carrying a current.

The nature of such a field is best shown by memorizing a simple rule known as the *right hand rule.* This rule says that if the thumb of the right hand indicates the direction in which the current is flowing the lines of magnetic force circle the conductor in the direction of the fingers of the closed hand. That is, if a wire carries a current as indicated in Fig. 16, an isolated north magnetic pole would move about the conductor in the direction of the arrows. While on the whole it is a poor policy to memorize several arbitrary rules such as the right hand rule, the dynamo and the motor rule, an exception may be made in the case of the right hand rule. It will be found that if this rule be remembered the interactions of wires

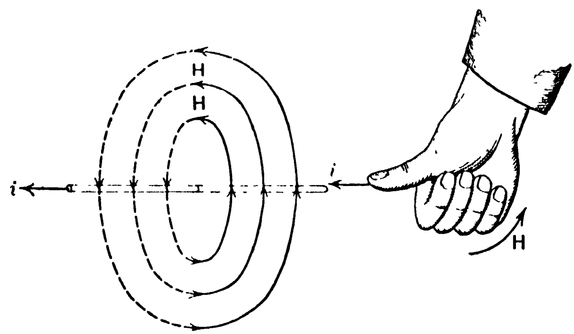


FIG. 16.—The Right Hand Rule.

carrying currents in magnetic fields of any sort can be predicted. (See Chapter XVII.)

19. AMPÈRE'S RULE AND DEFINITION OF UNIT CURRENT

In 1820, a brilliant experimentalist investigated the laws of currents. This was Ampère, and his study, which is the basis of all modern current measurements and is the essential to most current calculations, led to a general formulation of the field produced by the current. From these investigations we have the definition of the unit of electrical current and the foundation of the so-called *electromagnetic system of units.* Ampère showed that the force on a magnet pole in the neighborhood of a wire carrying a current is proportional to the length of the wire *taken perpendicular to the line joining the element to the point considered,* is proportional to the current, and is inversely proportional to the square of the distance from that current. The law may thus be expressed in the following equation: The force f on a magnetic pole m, at a distance r, from an element of conduction of length, ds, perpendicular to r, carrying a a current i_a, is given by

$$f = \frac{mi_a ds}{r^2}.$$

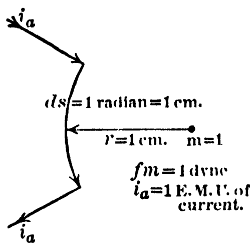


FIG. 17.—Definition of Unit Current.

If the conductor carrying a current be bent in the form of a circle of radius r, Fig. 17, the center of the circle will everywhere be equally distant from the wire, and the line from the center of the circle to the conductor will be everywhere perpendicular to the current. If the circle have a radius r the element ds will have a length $2\pi r$ cm. The force on a pole of strength m will then be

$$f = \frac{2\pi i_a m}{r}.$$

This at once leads us to a definition of electrical current in terms of the radius of the circle, the pole strength of the magnet, and the force exerted on the magnet. This definition of current is based on the magnetic action of the current, and furnishes the only *direct* means of establishing the magnitude of a current. Since the concept of a current at once establishes two other concepts based on the definition of the current, to wit, quantity and potential, it is seen that this definition constitutes the basis of a *system of units,* appropriately

named the *electromagnetic* system. The equation defining current is then the expression

$$i_a = \frac{r^2 f}{2\pi r m} = \frac{r^2 f}{m ds}.$$

The electromagnetic unit of current in the C.G.S. system will then be defined by the equation above if when $r = 1$ cm, $ds = 1$ cm normal to r, and m = unit magnetic pole, the force acting is 1 dyne. This leads to the usual definition given for unit electrical current; namely, *unit electromagnetic current is the current which flowing in a wire* 1 *cm long, bent into an arc of radius* 1 *cm, produces a magnetic field of* 1 *gauss at its center* (or exerts unit force on a unit pole at the center of the arc).

20. ALTERNATIVE DEFINITION OF UNIT CURRENT

The fact that the wire carrying such a current exerts a force on a unit magnetic pole at the center of the circle leads to another definition of unit current which will be of use later on. This definition is based on the fact that if the field produced by the current acts on a unit pole at the center with a force of 1 dyne, the pole acts on the wire producing the field with a force of 1 dyne. Now, a unit pole producing a unit field along the arc of a circle of 1 cm radius acts on a wire of unit length normal to that field with a force of 1 dyne, that is, the wire carrying unit current in a uniform magnetic field of unit strength experiences a force of 1 dyne. We can, therefore, also define *unit current as that current which flowing in a straight wire* 1 *cm long, placed perpendicular to a uniform magnetic field of unit strength, experiences a force of* 1 *dyne.* Symbolically, this may be stated as follows: Call i_a the current, l the length of conductor perpendicular to a uniform field H, Fig. 18, then the force in dynes on the conductor is given by $f = i_a l H$. Whence one can write $i_a = f/lH$. Therefore, i_a is one absolute electromagnetic unit of current if l is 1 cm, H is 1 gauss and f is 1 dyne.

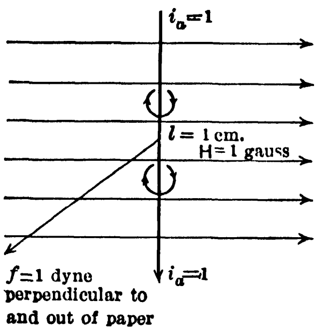


FIG. 18.—Alternative Definition of Unit Current.

In practice, the unit electrical current here defined is larger than the currents commonly dealt with. In order to have a *practical* unit which is convenient in magnitude (i.e., a unit of such magnitude

that values of the current must not always be written as a fractional part of the unit), *the practical unit of current is taken as one-tenth the absolute electromagnetic unit. It is called the ampere.* Frequently, the absolute electromagnetic unit of current is denoted by the term ab-ampere, meaning the absolute ampere, and will hereafter be denoted by the symbol i_a, while the ampere will be represented by the symbol i.

21. DEFINITION OF QUANTITY ON ELECTROMAGNETIC SYSTEM OF UNITS

The definition of current of electricity leads one at once to a definition of quantity of electricity in the electromagnetic system. The definition is based on an analogy between the flow of water and the flow of electrical current. The quantity of water which is flowing through a given pipe is defined in terms of the current multiplied by the time of flow. Similarly, the quantity of electricity is the current multiplied by the time during which it flows. *Thus the unit of quantity in the absolute electromagnetic system is that quantity which is represented by a current of one absolute electromagnetic unit flowing for one second.* Symbolically, this may be represented by writing quantity q is equal to current i_a times time t, or $q = i_a t$. The practical unit of quantity is again one-tenth the absolute unit of quantity. *It is represented by the flow of* 1 *ampere of current for one second. It is named the coulomb* and will be represented by the symbol Q.

22. APPLICATIONS OF AMPÈRE'S LAW

We now turn to applications of Ampère's rule to several simple cases which have practical uses.

(1) **The Law of Biot and Savart.**—This law gives us the field produced at any distance r from an infinitely long straight conductor carrying a current i_a. For practical purposes all that is needed is that the wire be straight and long compared to the distance r.

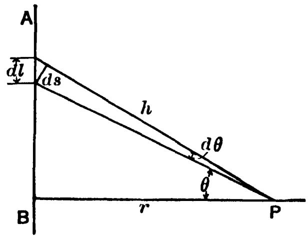


FIG. 19.—Field at a Point Due to an Infinitely Long Straight Conductor.

Let AB, Fig. 19, be a long straight wire. Let P be a point distant r cm from it. Let any element of the wire dl be chosen, the lower end of which when joined to P gives a line Pdl which makes an angle θ with the line r. Call $d\theta$ the angle subtended by dl

at P. Let ds be the projection of dl on the normal to the line joining dl and P, and let h represent the length of the line from P to dl.

Ampère's law says that contribution dH to the field at P by the projection of the element of wire dl on the normal to the line joining dl and the point P at which the field is required, is given by

$$dH = \frac{i_a dl}{h^2} \sin\, \theta = \frac{i_a ds}{h^2}.$$

Now,

$$ds = hd\theta,$$

hence,

$$dH = \frac{i_a d\theta}{h}.$$

As

$$h = \frac{r}{\cos\, \theta},$$

therefore,

$$dH = \frac{i_a \cos\, \theta}{r} d\theta.$$

To get the total field at P due to all the little elements, dl or their projections ds, one must integrate the contributions dH to this field for all the little elements dl, that is, one must integrate dH for elements dl running along the wire, from minus infinity to plus infinity distant from the perpendicular r. For the positive value of infinity, θ has the value $\frac{\pi}{2}$. For the negative value it has the value $-\frac{\pi}{2}$. Hence,

$$H = \int_{dH \text{ at } -\infty}^{dH \text{ at } +\infty} dH = \frac{i_a}{r}\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \cos\, \theta\, d\theta = \frac{2i_a}{r}.$$

The field, therefore, at any distance r from the long straight wire is merely twice the current in *absolute units* divided by the distance r. The field strength thus obtained will be expressed in gauss or dynes of force per unit pole.

There is a simple experimental proof of the correctness of this law. Let us place a bar magnet B on a frame F so that it is rigidly fixed perpendicular to a long wire W carrying a current, as in Fig. 20. The frame F is so mounted that the frame and bar magnet are free

to rotate about the long wire following the lines of force. Now suppose the north pole of the bar magnet be nearer the wire than the south pole. If the current i_a be flowing upward in the wire, the north pole will be urged in the direction of the arrows and the south pole will be urged in the opposite direction. Now the law of Biot and Savart says that

$$H_N = \frac{2i_a}{r_N}$$

and

$$H_S = \frac{2i_a}{r_S},$$

where H_N and H_S are the fields acting at the north and south poles, and r_N, r_S are the distances of the north and south poles from the wire. Therefore, it follows that $H_N r_N = 2i_a = r_S H_S$. Since the pole strength for the north and south poles are the same, the forces acting on the north and south pole should be given by the expression

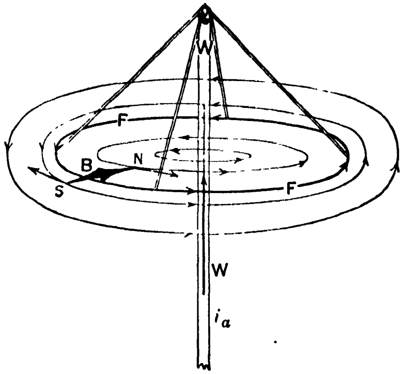


FIG. 20.—Experimental Proof of Biot and Savart's Law.

$$mH_N r_N = mH_S r_S.$$

But the forces mH_N and mH_S represent the actual forces acting on the two poles. Hence, the force moment causing the north pole to move in the direction of the arrow about the wire and the south pole to move in the opposite direction around the wire are given by multiplying these forces by the distances r_N and r_S from the wire. If the force moment on the north pole were greater than on the south pole the bar magnet as a whole would rotate around the wire. But the consequences of Biot and Savart's law above says that these force moments are equal. Consequently, if Biot and Savart's law is correct, there should be no motion. Experiment shows this to be true.

(2) The field at the center of a plane circular coil of radius r.—Ampère's rule says that the force on a unit pole by a current in a wire is given by

$$f = \frac{i_a ds}{r^2}.$$

In this case r is the radius of the coil, and ds is the length of the arc

normal to the line from the center of the circle to the coil. As the radius of the circular coil is always perpendicular to the arc, the expression for the *magnetic field* becomes

$$H = \int_0^H dH = \frac{i_a}{r^2}\int_0^{2\pi r} ds = \frac{i_a 2\pi r}{r^2} = \frac{2\pi i_a}{r}.$$

For n turns of wire in the coil, this will be n times as great, or $H_n = \frac{2\pi n i_a}{r}$.

(3) **The Field Produced by a Plane Circular Coil at a Point Distant d cm from the Plane of the Coil, Along the Normal to the Coil, at Its Center.**—Let the coil be represented by the circle drawn in perspective, in Fig. 21. Call O the center of the coil and let its axis be the line OB. A current i_a is flowing around this coil, the radius of which is r. It is desired to know the magnetic field in magnitude and direction at a point P, at a distance d from the center of the coil. Take any element ds of the coil and draw a line joining ds to P, and r joining it to O. According to the right hand rule the current flowing in the coil, as indicated by the arrows, produces a field at P at right angles to h, the line joining ds and P, as indicated by PK in the diagram. Let the length of the line from ds to P be h, and let θ be the angle between h and d. Now, it is obvious that if one regard all the elements ds in the circumference of the circular coil, there will be for each element an equal element ds', at the opposite end of the diameter of the circle. The fields produced by two such elements will be of the same magnitude but in directions such that they will in part neutralize each other; that is, their components normal to the axis will cancel each other. The resultant field will be the sum of the components of the fields produced by the elements ds which are not annihilated by the fact that each element ds is paired against a similar element at the end of its diameter. Therefore, the component of the field,

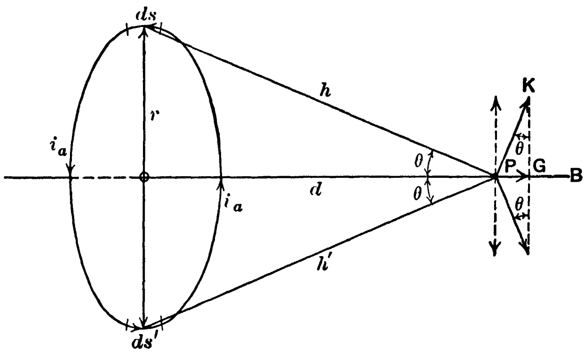


FIG. 21.—Field on the Axis of a Circular Coil.

due to the element ds which is of importance, will be the projection of the vector represented by PK on the line OP; that is, it will be the force due to the element ds multiplied by sin θ represented by PG. One is now in a position to apply Ampère's rule. This says that $dH = \frac{i_a ds}{r^2}$. In this case, the element of field in which we are interested is the component parallel to the axis of the coil and is consequently called:

$$dH_1 = dH \sin \theta.$$

Furthermore, r in Ampère's equation is the length of the line from ds to P, or h. If there are n turns of wire in the coil the expression for dH_1 must be multiplied by n.

Hence,

$$dH_1 = \frac{ni_a ds \sin \theta}{h^2}.$$

As

$$h = \frac{r}{\sin \theta}, \quad dH_1 = \frac{ni_a ds\, r}{h^3}.$$

But,

$$h^2 = r^2 + d^2,$$

and

$$H_1 = \int_0^{H_1} dH_1 = \frac{ni_a r}{(\sqrt{r^2 + d^2})^3}\int_0^{2\pi r} ds = \frac{2\pi n i_a r^2}{(r^2 + d^2)^{3/2}}.$$

23. THE TANGENT GALVANOMETER AND THE SINE GALVANOMETER.

In practice i_a is measured by means of the circular coil discussed in Case 2. The instrument formerly used for this is known as the *tangent galvanometer*. The term galvanometer indicates measurer of galvanic currents. The type of galvanometer receives its designation from the essential feature of the measuring device to be discussed. The *tangent galvanometer* makes use of the comparison of the field produced by the current, and the field of the earth. For this purpose it is necessary to know the earth's field accurately, and hence the necessity for the careful magnetometric measurement of the earth's field discussed in Chapter IV

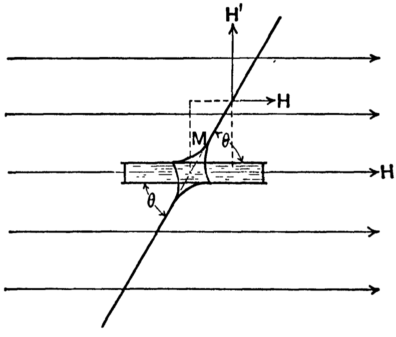


FIG. 22.—Principle of the Tangent Galvanometer.

Call the field of the earth H and represent it by the parallel lines of Fig. 22. Set the coil of the tangent galvanometer with its plane *parallel to the earth's field* as seen in the diagram. The field produced by this coil will then be at right angles to the earth's field. Beside the coil the tangent galvanometer has a compass needle mounted on an axis at the center of the coil. It is mounted to swing in a horizontal plane, and its angle with the plane of the coil can be read on a graduated scale. The magnetic needle of moment M indicated in the sketch will thus be subjected to torques from two sources. The field of the coil will cause it to try to set itself at right angles to the plane of the coil. The earth's field will cause it to try to set itself in the plane of the coil. As a result the needle will come to rest at an angle θ with the direction of the earth's field and the plane of the coil. This angle θ will be determined by the relative strengths of the earth's field H and the field of the current in the coil H', as the following consideration will show. The torque due to the field H is given by $G_H = HM \sin \theta$, where M is the moment of the needle (see page 74). The torque due to the current in the coil is $G_{H'}$, and is given by $G_{H'} = H'M \cos \theta$. At equilibrium $G_H = G_{H'}$, hence we can write $HM \sin \theta = H'M \cos \theta$.

Thus,

$$\frac{H'}{H} = \tan \theta.$$

Now H' is from Case 2 given by $H' = \dfrac{2\pi n i_a}{r}$.

Hence $\dfrac{2\pi n i_a}{rH} = \tan \theta$, and $i_a = \dfrac{rH}{2\pi n} \tan \theta$, where n is the number of turns of wire in the coil.

Thus as the earth's field is known and the constants of the tangent galvanometer are known, all that is necessary is to pass a current through the instrument, measure the angle θ, and i_a will be given in absolute units if the earth's field H is given in Gauss. The quantity $\dfrac{r}{2\pi n}$ is known as the constant of the galvanometer, for it is characteristic of the particular instrument. To get the current in amperes, the current as measured by the tangent galvanometer must be multiplied by the factor 10, so that $i = \dfrac{10rH}{2\pi n} \tan \theta$.

Another instrument which is analogous to the tangent galvanometer is the *sine galvanometer.* In this case, the coil that carries the

current is turned in the earth's magnetic field until the needle lies in the plane of the coil as shown in Fig. 23. In this case $G_H' = H'M$ since the field produced is perpendicular to the needle. G_H is, however, $MH \sin \theta$ as before.

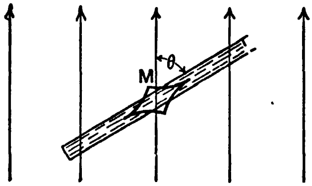


FIG. 23.—Principle of the Sine Galvanometer.

The equilibrium thus yields,

$$H'M = MH \sin \theta,$$

and, therefore,

$$i_a = \frac{Hr}{2\pi n} \sin \theta.$$

This instrument is not used at present but furnishes an excellent illustration of the principles involved.

CHAPTER VI

POTENTIAL DIFFERENCE AND WORK IN AN ELECTRIC CIRCUIT

24. THE CONCEPT OF POTENTIAL

WHEN we regard the flow of electricity from a voltaic pile through a long fine wire, we note that to increase the current it is necessary to connect the piles up in series, that is, each zinc element of a pile must be connected to the copper element of the next pile, and so on for the number of piles used, as illustrated in Fig. 24. As the cells are added the current through a given wire to which the end cells are connected will increase nearly in proportion to the number of cells. This increase in flow of electricity (for so current must be considered as a result of the study of static electricity) set up by the increase in the number of cells might be regarded as an increased flow following from a sort of increased electrical pressure. Already as early as 1734 the fact that like kinds of electricity repel each other and unlike kinds attract each other as well as that electricity could travel along a wire, had led to the notion that electricity was a fluid. In fact, Benjamin Franklin had proposed the theory that there was a weightless electrical fluid, or positive electricity, which repelled itself and could flow especially well over metals. A neutral body with an *excess* of positive electricity on this concept is *positively charged* and a neutral body having some of its electrical fluid removed has *lost* its *positive electricity*, or has become *negatively charged.* The positive electricity was identified with the so-called vitreous electricity obtained by rubbing glass or glass-like bodies with silk. Thus, as a result of attractions between unlike charges and repulsions between like charges, accumulations of positive electricity will always, if given a chance, flow to places where the density of positive electricity becomes less. In other words, if electricity is likened to a fluid it will flow until the electrical pressure due to self-repulsion is every-

FIG. 24.—Voltaic Piles in Series.

where equal. Thus we see that on this view it must be the electrical pressure, or rather the analogue to fluid pressure in electricity, which is increased by increasing the voltaic piles in series and which determines the current.

Since the analogue to fluid pressure in electrical circuits determines the flow of electricity we must be able to define this factor in current flow clearly and to understand its properties. It is, therefore, necessary to study the concept of the electrical potential, for so this analogue is called, in order to understand electrical circuits. To accomplish this easily it is best to consider first the behavior of water, a true fluid, whose properties and behavior we can easily visualize. Once an understanding in terms of a fluid has been gained the transfer to the electrical case is a comparatively simple matter.

When we have water in a tank A, Fig. 25, which can be connected to a tank B by means of a valve V, we know that if the level in the tank A is above that in B the water will flow from A into B, so that when equilibrium is reached the levels in the two tanks are equal. The flow follows from the fact that the pressure at a point depends on the depth of water above the point, and for points at the same level in A and B the pressure must be equal to prevent flow.

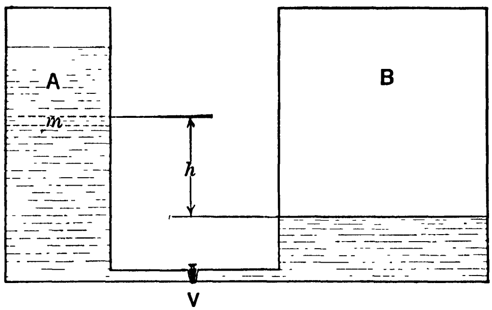


FIG. 25.

Hence, the difference in pressure due to differences in level causes flow. Now regard a small mass dm of water in tank A at a level h cm above its final level in tank B after flow has ceased. When it moves down to the level in the tank B, i.e., through a distance h, it does work equal to $dmgh$, for to begin with it had potential energy $dmgh$ above its final level in the tank B. That is, owing to the force of gravity, dmg on the mass dm when it moves a distance h an amount of work $dmgh$ is done on it. This work came from its energy of position or potential energy in the earth's gravitational field. If the whole mass m of water in the tank had moved through the average distance h, then again work equal to mgh would have been done. One may now inquire where the energy went, for at rest at the new level there is no energy obvious. Had tank B been absent and the valve V and the tube had a large diameter, the water would have emerged from A

in a stream as shown in Fig. 26*a*. This stream could have been used to turn a paddle wheel and would have given work. The stream of water thus possesses ability to do work in virtue of its motion; that is, it has energy of motion, or kinetic energy $\frac{1}{2}mv^2$. If the pipe and valve V had been very small the water, as shown in Fig. 26*b*, would merely have trickled out. There would have been no velocity, and thus no energy of motion, and the potential energy mgh would have been consumed in the narrow tube and valve in overcoming friction. Such energy consumption as Joule showed goes into heat.

Thus it follows that when as a result of pressure differences water flows work is done which may be converted into heat as friction or into energy of motion in turning a paddle wheel. The work available for these processes whether it goes entirely to heat or only partly to

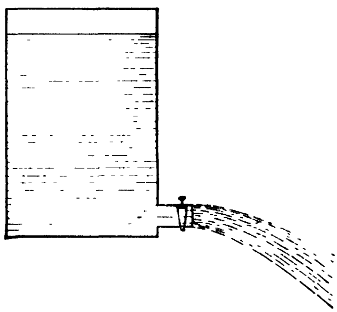


FIG. 26*a*.

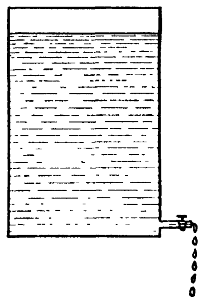


FIG. 26*b*.

work and partly to heat is the potential energy mgh. Now the mass of water m is the mass or quantity of water which moves, and gh represents what is known as the change in gravitational potential or level involved in the process. We can thus write gh = P.D. where P.D. signifies potential difference. It is seen that g is the characteristic constant of the gravitational field, and multiplied by the change in level it presents the change in potential or potential difference. Since P.D. = gh, and work $w = mgh$ for a mass m transferred, we see at once that we can define P.D. by the expression, P.D. $= gh = \frac{w}{m}$. Thus P.D. or potential difference is simply the work done per unit quantity or mass transferred. This definition says nothing of the nature of the process, it simply gives a new quantity, potential difference, which is proportional to the head or pressure that drives the water through the tubes, and is defined as work

per unit quantity transferred, i.e., in terms of the performance observed. Since the current depends on the head or potential difference we see that this new concept is fundamental to an analysis of currents in circuits.

Now in electricity an exactly analogous situation exists. By putting more galvanic cells in series apparently the electrical pressure is raised as more current flows. We may then expect that the change of electrical level as a result of flow of electricity will be accompanied by phenomena similar to those which occur with the change in level of water and its flow. In fact if we connect a voltaic pile across a motor it turns over, and mechanical work is done. We can also connect the pile to electrodes immersed in water resulting in the expenditure of chemical work in decomposing water. Finally when we connect the terminals together by a wire which has in analogy to the narrow tube with water a high resistance, the energy of the current will go into friction and the wire will be observed to rise in temperature as electrical current flows in it.

We may then without further discussion consider that the electrical difference in level which causes flow of current or electrical potential difference P.D. can in analogy to water be defined by an equation P.D. $= \frac{w}{q}$, where as before we represent by w the work done when the quantity of electricity, q (equivalent to the mass of water m in our water analogy), has been transported through the difference in potential P.D. *The potential difference between two points in a circuit which represents the difference in electrical level causing the flow of current between the points is then defined as the work to transport unit quantity of electricity from the one point to the other.*

Now since we have adopted Franklin's terminology of the nature of electricity and arbitrarily *assumed* that vitreous electricity is the *positive* or real electrical fluid, we must *assume* a convention of flow that conforms to this terminology and to the mutual repulsion of like kinds of electricity. We therefore say that a current of positive electricity flows from a positive potential to a lower positive potential or to a negative potential. The reference point for electrical pressure or potential is the potential of the earth, which is arbitrarily for convenience chosen as 0. As a rule in current electricity we are not concerned with anything but the potential differences between points in a circuit. Where a current is flowing we must assume, as in a tube with water flowing through it, that there is a fall in potential or electrical level along the resistance or conduit carrying the current and this can be demonstrated experimentally as will be seen. Since we

have defined electrical potential difference as P.D. $= \frac{w}{q}$, and since q by the last chapter is defined as $i_a t$, therefore we can write P.D. $= \frac{w}{i_a t}$, or $w =$ P.D. $i_a t = q$P.D.

25. THE DEFINITION OF POTENTIAL IN THE ABSOLUTE ELECTROMAGNETIC SYSTEM

The question then arises: how does the energy which is consumed when electricity falls from the potential of the positive pole to that of the negative pole in flowing through a wire manifest itself? Since the only work that is done in flowing through the wire is to overcome the resistance of the wire to the flow, in an analogy to the case of water, we expect the frictional energy which is consumed in the wire to manifest itself in the production of heat. The truth of this assertion can be readily seen by connecting a wire across an electrical power main and observing the resulting rise in temperature of the wire. It is consequently very simple to relate the electrical potential difference between two points on a wire and the quantity of electricity which flows through the wire, to the work done, for we can measure the work in terms of the heat quantities liberated. The equation above P.D. $\times\, i_a t = W$ may now be equated to HJ where H is the heat liberated in calories and J is the mechanical equivalent of heat relating the calorie to ergs of work. Thus we can write our equation

$$\text{P.D.} \times i_a t = W = HJ.$$

This at once leads to the definition: *Potential difference between two points in a circuit is the work done when unit quantity of electricity is transferred from one point to the other.*

Symbolically, this may be seen from the equation above,

$$\text{P.D.} = \frac{W}{q} = \frac{W}{i_a t} = \frac{JH}{i_a t},$$

hence if $q = 1$ unit of electromagnetic quantity and W is 1 unit of work, P.D. is 1 electromagnetic unit of potential.

According to this definition, the P.D. between two points is unity in the absolute electromagnetic system when a current of one absolute electromagnetic unit, flowing for one second, does one erg of work, or when one erg of work must be used to move unit quantity of electricity, in the absolute electromagnetic system, from one point to the other, against the potential existing.

Now, an absolute electromagnetic unit of quantity is a very large quantity of electricity. If we are to do as small an amount of work as an erg when we move this from one point in the circuit to the other, the absolute unit of potential difference must be very small. Another way of arriving at the same conclusion comes from the heat relation: P.D. $= \dfrac{JH}{i_a t}$. Now if P.D. is unity, H is $i_a t/J$. Since for unit quantity $i_a t$ equals unity, H will be epual to $\dfrac{1}{4.18 \times 10^7}$ calories.

This quantity of heat is too small to measure practically. For ordinary purposes, the practical unit of potential difference is a larger unit called the *volt. The practical unit of potential difference is the potential difference existing when it takes* 10^8 *ergs of work to move one absolute electromagnetic unit of quantity against the existing electrical field.* Denoting this practical unit of potential difference, the volt, by V, we have $V = 10^8 \times \text{P.D.}_{\text{E.M.U.}}$ Thus, if we use volts,

$$H = 10^8 \frac{\text{P.D.} \times i_a t}{4.18 \times 10^7} = 2.4\ V i_a t.$$

Therefore, H (in calories) $= 2.4 \times V$ (in volts) $\times\ i_a$ (in absolute units) $\times\ t$ (in seconds). Remembering that an absolute unit of current is 10 amperes the expression becomes, H (in calories) $= 0.24\,V$ (in volts) $\times\ i$ (in amperes) $\times\ t$ (in seconds). Thus, to measure the potential difference in the electromagnetic system all we need do is to pass the current through a wire, measure the current in amperes, the time in seconds, and the heat developed in calories. From the above equation, the heat liberated will then give the potential in volts by means of the equation

$$V = \frac{H}{0.24 it}.$$

What has gone before gives the definition of potential in the absolute electromagnetic and in the practical system of units. The value of the potential difference as above defined can be measured by the arrangement shown in Fig. 27, which is the fundamental means of determining the value of the volt in the electromagnetic system. C is a calorimeter in which a wire coil of resistance R is placed. The coil is surrounded by water so that the weight of water plus the water equivalent of the calorimeter furnish a convenient mass for determining the heat liberated in the wire. The source of potential which is to be measured is connected to the two wires of low resistance of the coil R coming from the calorim-

eter C. In series with this coil is placed a galvanometer or ammeter A of low resistance for measuring the current. By taking the reading of the current on the ammeter with the potential turned on for a known time t, the potential can be evaluated at once by the equation above

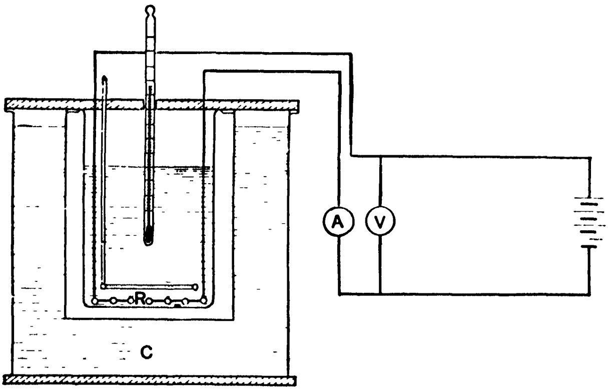


FIG. 27.—Calorimetric Measurement of Potential.

from the heat given to the calorimeter and the reading of the voltmeter V can be checked.

26. THE DEFINITION OF ELECTRICAL POWER CONSUMPTION

From the equation developed that P.D. $i_a t = W$, we get the relation that

$$\text{P.D.} i_a = \frac{W}{t}.$$

$\frac{W}{t}$ is the time rate of doing work. That is, it gives the *power* developed. *Thus, P.D. (in absolute E.M.U.)* $\times$ *current (in absolute E.M.U.) gives the power in ergs per second.*

V. (in volts) $\times$ *i (in amperes) gives power in* 10^7 *ergs per second.*

This work is a joule per second, where a joule *equals* 10^7 *ergs, and the unit, joule per second, is called the watt.*

Therefore one may write: volts times amperes equals watts.

The watt is the practical unit of power in the C.G.S. system and is of great importance in engineering practice. For large power output the kilowatt = 1000 watts is used.

The relation that volts times amperes times seconds equals 10^7 *ergs* leads to the interesting relation that volts times amperes times seconds equals volts times coulombs, or equals joules of energy. It

is worth while to remember these relations, as they are of considerable importance in the solution of problems.

27. THE CONCEPT AND DEFINITION OF ELECTRICAL RESISTANCE

We now have two absolute units of electricty defined in terms of the electrical current. These are the current i_a and the potential difference P.D. We have also related them to the practical units – the ampere and the volt. It is found *experimentally* that the current in a given circuit is proportional to the voltage or potential difference applied. That is, for a given circuit,

$$\frac{\text{P.D.}}{i_a} = \text{a constant.}$$

If we keep P.D. constant as we go from one circuit to another, we find that i_a varies with the form and dimensions of the circuit. That is, the ratio $\frac{\text{P.D.}}{i_a}$ depends on the form and dimensions of the circuit, and is a *constant of the circuit*. This constant is called the *resistance* of the circuit. We consequently define the *resistance of a circuit, or of a portion of a circuit, as the ratio of the potential difference to the current produced*. Thus, the unit of resistance in the absolute electromagnetic system is the resistance of a circuit which allows a P.D. of one absolute electromagnetic unit to maintain a current of one absolute electromagnetic unit in the circuit. Since the absolute unit of P.D. is small while the absolute unit of current is large, this absolute unit of resistance is *very* small. *In practice, the unit of resistance used is the resistance which permits a potential difference of* 1 *volt to maintain a current of* 1 *ampere through it. This unit is known as the ohm.* Symbolically, this may be represented by

$$\frac{V \text{ (in volts)}}{i \text{ (in amperes)}} = R \text{ (in ohms)}.$$

Since the volt is 10^8 absolute E.M.U. and the ampere is 10^{-1} E.M.U., the absolute E.M.U. of resistance is 10^{-9} ohms.

This law is known as Ohm's law in honor of G. S. Ohm, who was the first to deduce this relation between potential difference and current, and to show the significance of potential difference. He derived the law using the analogy between the flow of heat in a circuit and the flow of electricty in a circuit. This was in 1827, shortly after the great mathematician Fourier, in 1822, worked out the laws of the flow of heat.

It is important to notice in this connection that we now have a ratio of two quantities defined in the absolute electromagnetic system, potential difference and current. From the ratio of these two we find that a given circuit has a constant ratio, and *that this is characteristic of the form of the circuit. We, therefore, use the ratio of potential difference and current to define a third quantity which is a characteristic of the circuit, called resistance.* The two quantities potential difference and current, are *fundamental quantities,* since they are directly derived by measurements given for quantities in terms of the absolute C.G.S. system. Resistance, however, is a *derived* unit, for it is derived from a ratio of two fundamental units.

It is essential to avoid the vicious cycle so often indulged in by elementary students of physics of defining resistance in terms of current and potential and then turning around and defining potential in terms of current and resistance. However, we may define the quantities, two of them must always be fundamental and one of them derived. It is more logical, in the treatment to follow, to define potential and current as fundamental units because of their relation to heat and magnetic fields and to treat resistance as a derived unit than any other arrangement. (See the Introduction.)

28. JOULE'S LAW OF HEATING

We can now briefly apply the relation of Ohm's law to the heating effect of an electric current.

For we wrote,

$$H = 0.24\ Vit.$$

Since

$$\frac{V}{i} = R,$$

we can write

$$H = 0.24\, i^2Rt.$$

This says that the heating effect is proportional to the square of the current, and to the resistance of the circuit. The heating effect varies from circuit to circuit as the resistance varies. This law is known as Joule's law of heating.

A nice verification of Joule's law can be seen in the experiment to be described. Three coils, R_1, R_2 and R_3, of exactly equal resistance R are wound on insulating frames as shown in Fig. 28. The resistances R_1 and R_2 are connected together and the leads from a potential main are connected across them directly, one of the wires

from the main, however, first passing through the third coil R_3. The current i from the main flows through R_3 and then splits, one-half going through R_1 and one-half going through R_2. Now coils R_1 and R_2 are immersed in separate beakers with 200 cm³ of water in each

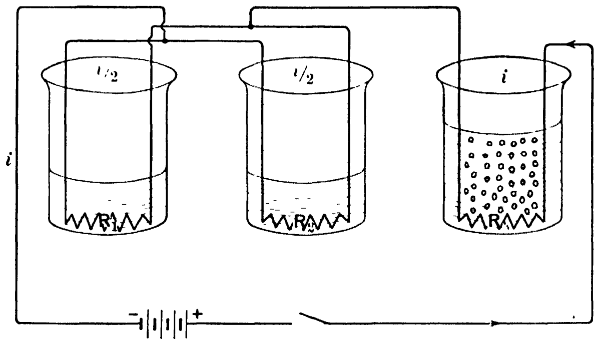


FIG. 28.—Verification of Joule's Law of Heating.

of them while R_3 is in a beaker with 400 cm³ of water. The heat in the beakers with R_1 and R_2 is $H_1 = 0.24\, R\left(\frac{i}{2}\right)^2 t$ while that in the beaker with R_3 is $H_3 = 0.24\, Ri^2t$. Since the heat, H_3, in R_3 is four times H_1, that in R_1 and R_2, then the 400 cm³ of water in the beaker with R_3 will boil long before that in the beakers R_1 and R_2 with 200 cm³ of water each.

CHAPTER VII

ELECTRIC CURRENTS AND RESISTANCE

IN the last chapter it was shown that in each circuit there is a quantity defined by the ratio of potential difference to the current flowing which was called the resistance of the circuit. This *general law,*

$$\frac{V}{i} = R,$$

holds for any part of the circuit as well as for the circuit as a whole, provided i is the total current flowing between the points across which the potential V exists. If, by chance, R is known, and one can measure i, the potential is given at once by the expression $V = iR$. This potential difference, defined by iR, is often spoken of colloquially as the iR drop in the circuit, as it measures the drop in potential between the two points delimiting R.

The potential difference between any point on a wire and the point of lowest potential on the wire drops continuously as one approaches the point of lowest potential. An analogy to the case of flow of water in a long tube where the hydrostatic pressure, or the potential, at the orifice is 0, while at the other end of the tube it has the maximum potential applied, shows that if a series of manometers are placed at various points along the tube they indicate a progressive fall of potential, or pressure, down to the open orifice.

29. NATURE OF RESISTANCE AND ITS VARIATION

How the potential varies along a conductor can be calculated from the relation just given, $V = iR$, if we know i and the value of R over the portion of the circuit considered. It therefore becomes necessary to study the laws of resistance. *Experiment* has shown that for a given uniform conductor, the resistance R is proportional to the length l, and is inversely proportional to the area of cross-

section A. The relation between these three quantities is given by the expression

$$R = \frac{R_S l}{A}.$$

In this equation R_S is the constant of proportionality and is a characteristic constant of the material of the conductor. Actually, it is the value of R when $l = 1$ cm, and $A = 1$ sq. cm. Thus R_S is the resistance, in ohms, of a unit cube of the material. As R_S is a characteristic of the material, *it is called the specific resistance* of the material. The value of R_S for any substance may be found in physical tables. The reciprocal of R_S, that is,

$$\frac{1}{R_S},$$

is called the *conductivity*. The conductivity of known substances ranges from that of exceedingly good conductors (specific resistance 2×10^{-6} ohms), down to conductors in which the passage of current can be observed only by the most careful measurements (specific resistance 4×10^{15} ohms). Among the best conductors are the metals. These, headed by silver and copper, passing down through metals such as iron, bismuth, etc., and through the alloys, constitute a class by themselves, the nature of whose conductivity will be discussed more in detail later. Between the metallic conductors and the other class of conductors there is a rather large gap. The second class of conductors begins with solutions and some fused salts, and goes down through various non-metallic solids to substances like stone, rubber, paraffine and sulfur. The solutions are fairly good conductors, while sulfur is perhaps the best non-conductor which is known to us. It might be stated that the difference between the two classes of conductors mentioned is that the metallic conductors act by transporting the current by means of the minute particles of negative electricity, the electrons. These are so small in size that they meet with comparatively little resistance in moving through the space lattice occupied by the atoms. Consequently, there is comparatively little resistance to the flow of current. On the other hand, substances like solutions and fused salts conduct the electricity through the agency of charged atoms, or molecules, of the substances in them. Owing to the size of these atomic carriers, compared to electrons, the conductivity of the second type is lower than that of the electronic type.

30. RESISTANCE AND TEMPERATURE

The specific resistance, R_S, while it is a constant characteristic of the particular metal, and while it is changed very greatly by the minutest traces of impurity, is also a function of temperature. In all metallic conductors, the resistance increases as the temperature increases. In general, the resistance R_T at any temperature T in degrees C is given by $R_T = R_S(1 + aT)$. Here a is called the *temperature coefficient of resistance.* For most substances a is small, and the equation given holds through rather wide ranges. A more accurate expression for the resistance as a function of the temperature is given by an empirical equation of the form

$$R_T = R_S(1 + aT + bT^2).$$

With the very accurate devices for measuring potentials, or resistances, which exist today, it is possible to make exceedingly accurate measurements of the resistance of wires at various fixed points on the temperature scale. By plotting the curve through these points it is possible to have a very accurate temperature scale in terms of the resistance of the material, and the *resistance thermometer* today is the basis of most refined heat measurements over limited ranges. While, in general, the resistance decreases as the temperature decreases below the freezing point of water, the slope of the curve is such that the resistance appears to approach a finite value at the lowest temperatures. In certain substances, however, notably lead, the curve between resistance and temperature on the absolute scale takes a sudden very definite turn to much lower values, at very low temperatures. In the laboratory of Kammerlingh-Onnes, in Leyden, Holland, the conductivity of lead at the temperature of liquid helium (about 4° absolute) has been measured. At these temperatures the resistance of lead is very nearly 0 and an electric current generated by means of induction in a small leaden circuit continues to flow for periods as long as days before the frictional loss due to the resistance consumes the energy. This state of extreme conductivity at low temperatures is known as *super-conductivity.* The exact nature of this super-conductivity has not yet been determined. It is probable that at these low temperatures the atoms in the lead crystal arrange themselves in such a fashion that there are almost forceless tubes or channels through which the electron can flow, so that when once set in motion the electron continues its motion, losing no energy to set the atoms in vibration. The phenomenon, however, is limited to a very few of the elements only (Pb. 7.3°, Ta 4.5°, Hg 4.2°, Sn 3.7°,

In 3.4°, Tl 2.5°, Th 1.4°, Au-Bi alloys 2.15°; temperatures all in degrees absolute).

31. COMBINATION OF RESISTANCES IN SERIES AND PARALLEL

Having defined specific resistance we may now turn to the question of the laws of resistances in circuits.

Case 1. Law for resistances in series.—Recalling Ohm's law, $V = iR$, and remembering that it applies to the parts of the circuit as well as to the whole circuit, we can write for the circuit pictured in Fig. 29 that

$$V = V_1 + V_2 + V_3,$$

and hence

$$iR = iR_1 + iR_2 + iR_3,$$

where R_1 refers to V_1, R_2 to V_2, and R_3 to V_3. Therefore, we can

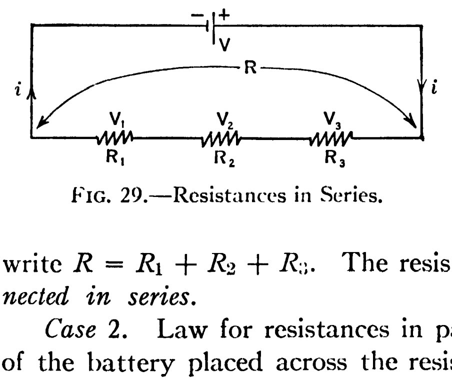


FIG. 29.—Resistances in Series.

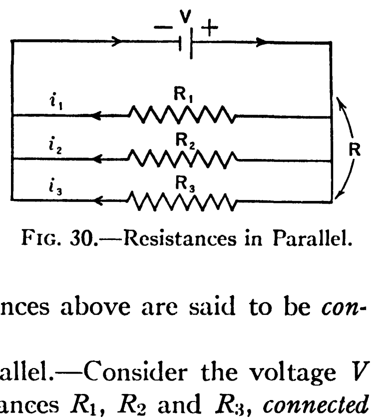


FIG. 30.—Resistances in Parallel.

write $R = R_1 + R_2 + R_3$. The resistances above are said to be *connected in series.*

Case 2. Law for resistances in parallel.—Consider the voltage V of the battery placed across the resistances R_1, R_2 and R_3, *connected in parallel* as shown in Fig. 30. The currents through R_1, R_2 and R_3 are, respectively, i_1, i_2 and i_3. The potential across them is the same, namely, V. From Ohm's law, we have

$$\frac{V}{R} = i, \quad \frac{V_1}{R_1} = i_1, \quad \frac{V_2}{R_2} = i_2, \quad \frac{V_3}{R_3} = i_3.$$

Since

$$i = i_1 + i_2 + i_3,$$

therefore,

$$\frac{V}{R} = \frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3}.$$

But

$$V = V_1 = V_2 = V_3,$$

whence

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}.$$

The resistances here are said to be *connected in parallel.*

As an example of the application of these laws one may take the simple case depicted in Fig. 31, namely, the combination of the resistances R_1, R_2 and R_3. It is required to find the total resistance R. From the laws of resistance above we can write:

$$R = R_1 + R_{23} \text{ (resistance of } R_3 \text{ and } R_2 \text{ in combination).}$$

Since

$$\frac{1}{R_{23}} = \frac{1}{R_2} + \frac{1}{R_3} = \frac{R_3 + R_2}{R_2 R_3},$$

therefore,

$$R = R_1 + \frac{R_2 R_3}{R_2 + R_3}.$$

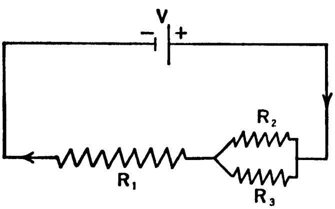


FIG. 31.—Series-Parallel Combination of Resistances.

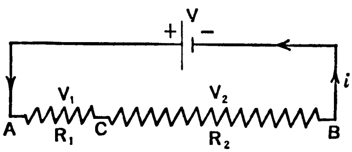


FIG. 32.—Fall of Potential along a Wire.

Case 3. Fall of Potential along a Wire.—Another case to which this analysis may be applied is the *case of the fall of potential along a wire.* Given the wire represented in Fig. 32, whose total resistance is R, and assume that there is a potential difference V across it. There is flowing in the circuit a current i. Let us call the positive end of the wire A and the negative end B. It is required to determine the fall of potential between A and some point C, and between C and B. Call R_1 the resistance in the wire from A to C and R_2 the resistance in the wire from C to B. Call the potential difference sought from A to C, V_1, and that from C to B, V_2. Now, by Ohm's law, V the total potential equals R times i. Since the same current flows in the parts AC and CB one has $V_1 = R_1 i$, $V_2 = R_2 i$, and $V = Ri$.

Since

$$R = R_1 + R_2,$$

therefore,

$$V = V_1 + V_2.$$

From this we get at once that

$$\frac{V_1}{V} = \frac{V_1}{V_1 + V_2} = \frac{iR_1}{iR_1 + iR_2} = \frac{R_1}{R}.$$

In a similar fashion,

$$\frac{V_2}{V} = \frac{V_2}{V_1 + V_2} = \frac{R_2 i}{R_1 i + R_2 i} = \frac{R_2}{R}.$$

Finally,

$$\frac{V_1}{V_2} = \frac{R_1 i}{R_2 i} = \frac{R_1}{R_2}.$$

This simply states that the fall of potential across AC or CB is to the potential across AB as the resistance of AC or CB is to the total resistance, and further that the fall of potential across AC is to the fall of potential across CB as the resistance of AC is to the resistance of CB. Since it is possible accurately to compare resistances or to measure them, *it is at once possible by properly choosing the resistances to obtain any fraction of a given fall of potential.* This is the principle of the *potential divider,* or *potentiometer.* It is a principle underlying so many electrical measurements and comparisons that it is perhaps one of the most important principles of current electricity.

32. POTENTIAL AND ELECTROMOTIVE FORCES IN CIRCUITS

The fact that Ohm's law holds for the whole of the circuit, as well as for parts of the circuit, leads to a discussion of considerable importance in the treatment of generators of electricity. Consider a battery and let it be connected by wires to a galvanometer, as indicated in Fig. 33.

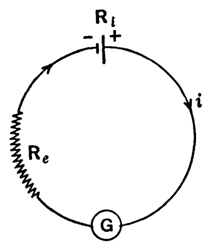


FIG. 33.—External and Internal Resistance in a Circuit.

It will later be seen that any generator, either battery or dynamo, will produce a *certain maximum potential difference between its terminals.* Thus depending on the chemical nature of the constituents a battery will produce a given *electromotive force or maximum potential difference* whose value can be calculated from the knowledge of the energy involved in the chemical transformations taking place, as we shall see in a subsequent chapter. Again, a dynamo will also produce a certain maximum potential difference between its terminals, which depends purely on the design of the dynamo and the speed

with which its armature is being rotated. *This maximum potential is called the electromotive force and is represented by the symbols E.M.F.* In both cases the E.M.F. obtained can be determined directly only by a static measurement or one involving no flow of current. The reason for this will become evident immediately.

Every battery as well as every dynamo has a resistance within itself (i.e., resistance of the electrolyte or of the armature windings). This is called the *internal resistance.* If we take current from a battery or from a dynamo, in order to maintain the potential difference at the terminals, the electricity which is being used must be supplied as rapidly as it is being drawn away. In order to accomplish this two conditions must be met. (*a*) The supply of the current to the external circuit must be kept constant. (*b*) The current through the generator to the terminal must be kept at the same value as in the external circuit. Thus a current must be forced through the internal portions of the generator to charge the plates at the same rate at which the plates are discharging to the external circuit. The current is being fed to the external circuit at a certain potential V which, by Ohm's law, is equal to the external circuit resistance R_e times i the external current. If we call R_i the internal resistance the current i flowing through it to supply the external current i requires that a fall of potential iR_i exist across the interior of the generator. Hence, if the current i is to be maintained the total potential, or the electromotive force (generally written E.M.F.) E, developed by the generator, must be equal to the sum of the iR_i drop inside the generator and the iR_e drop in the external circuit. Hence we must write

$$E = iR_i + iR_e = iR_i + V.$$

It is therefore seen that *to maintain a potential V in an outside circuit the generator supplying the current must produce an electromotive force E which is greater than V by the iR_i drop in its internal resistance.* Only when $i = 0$, so that $iR_i = 0$, can the potential V, maintained by a generator in the external circuit, be equal to the electromotive force. This means that even if we use a high resistance voltmeter with which to read the potential of a generator of some sort, the potential read will be the external potential drop through the coils of the voltmeter, and it will be less than the E.M.F. of the generator because the voltmeter draws a current even though it be a small one. It is seen that the difference between the electromotive force and the external potential maintained is merely the product of the current and the internal resistance or the internal drop of potential. If a method is capable of measuring potential with zero current the

method will give the E.M.F. of a generator. Such measurements may be made by means of *static* voltmeters, quadrant electrometers, or by means of the potentiometer. (See Chapter IX.)

From all this it may be seen that essentially the term potential drop or potential difference applies in the case of current flow to the difference of potential which maintains a current i through a resistance R in any portion of a circuit. *The electromotive force or E.M.F. is the total potential which is required to be generated by a source of current in order to maintain the current flow.* The E.M.F. is usually a function of the chemical constitution of the cell or of the magnetic conditions and speed of a dynamo. The electromotive force can be used to calculate the current through an external circuit at a given external potential provided the internal resistance is known, or vice versa. While the E.M.F. can only be measured by 0 current methods directly it can be calculated, and R_i can often be measured directly. If neither R_i nor E is known we can, by using two values of R_e and observing the two values of V maintained across R_e, obtain two equations of the form below which may be solved simultaneously for R_i and E. From the equations $E = i(R_e + R_i)$ and $V = i(R_e)$ we can at once write that $\frac{E}{V} = \frac{R_e + R_i}{R_e}$. Accordingly, if by varying R_e two values of V (V and V' corresponding to R_e and R_e') are obtained, we have the relations

$$\frac{E}{V} = \frac{R_e + R_i}{R_e}, \text{ and } \frac{E}{V'} = \frac{R_e' + R_i}{R_e'},$$

from which E and R_i can at once be found.

That the phenomenon discussed above is an important one may be seen from the following example: Assume a battery having an internal resistance $R_i = 2$ ohms, and an external resistance $R_e = 4$ ohms, and assume that the electromotive force E is 6 volts; then the potential V across the 4 ohm resistance would be $\frac{V}{E} = \frac{4}{2+4}$. Thus $V = \frac{2}{3} E = 4$ volts. It is thus seen that in working with small resistances we must be careful to distinguish between the electromotive force of a cell and the potential difference given by it. An ordinary voltmeter has a resistance of the order of 1000 ohms. Consequently, if placed across 5 dry cells with perhaps a total internal resistance of 10 ohms, the potential read by the voltmeter is less than the electromotive force of the cell in the ratio of $\frac{1000}{1010}$. That is, it is 1 per cent in error. For very high resistances where R_e is very much greater than R_i, V can be said to approach E.

CHAPTER VIII

GENERALIZATION OF OHM'S LAW

33. KIRCHHOFF'S FIRST LAW OF DIVIDED CIRCUITS

WHILE Ohm's law is applicable to a great many of the simpler circuits, it is not sufficient for determining the currents in the case of complicated circuits. There are many problems, such as the problem of the Wheatstone bridge when it is not in balance, in which the existing currents and potential drops cannot be found by the simple Ohm's law. The reason is that Ohm's law is a special case of a far more general set of relations from which enough equations may be obtained to solve the problem. The method of treatment which we are going to take up is that due to Kirchhoff and originated about 1842. It is indispensable in the practical study of any circuits, and electrical engineers cannot work conveniently unless they have a good command of the use of these laws. The laws themselves are simple enough. The method of applying them is, however, more difficult. In what follows the laws will first be stated and will be applied to a simple case, namely: the Wheatstone bridge. In a second instance, a numerical problem will be worked out showing the actual method of approach to the study of any problem.

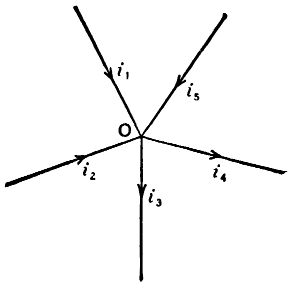


FIG. 34.—Current Flow at a Point in a Complex Circuit.

The first of Kirchhoff's laws says this: *The sum of all currents flowing into any point in a circuit must be* 0 *if taken with due regard to sign.*

Regard, for example, the 5 wires radiating from the point O which may be a binding post in any portion of a complex circuit, Fig. 34. There are flowing towards or from this point, 5 currents, represented respectively by the letters, i_1, i_2, i_3, i_4, and i_5. The arrows in the diagram indicate the direction in which the currents are assumed to flow. Now Kirchhoff's law merely states that, since the currents which are flowing towards this point O can only flow along the wires

and since there is no accumulation of electricity at the point, which will be the case *when a steady state exists,* the sum of the currents flowing *towards the point* must equal the currents flowing *away from that point.* Of course, during the first instant when the currents start to flow, as the point O is being raised to its equilibrium potential, there will be an accumulation of electricity. This, however, is a *transient* phenomenon and we are dealing at present only with equilibrium conditions in a circuit. Kirchhoff's laws for the particular circuit illustrated would then be expressed by the equation

$$\Sigma i = i_1 + i_2 + i_3 + i_4 + i_5 = 0.$$

The i's must be taken with due regard to sign.

34. KIRCHHOFF'S SECOND LAW OF DIVIDED CIRCUITS

Kirchhoff's second law is a more general statement of Ohm's law. In this case we are concerned not with a point in a circuit but with what we shall term a mesh. A mesh is any continuous circuit which may be a portion of a much more complex system. Thus, in Fig. 35, we may term the triangles ABC, and BCD, and the quadrilateral $ABDC$ meshes. The essential feature of the mesh is that there is a continuous circuit starting from a point and coming back to that point. This circuit may contain galvanometers, batteries, or current generators. The second law says that *in such a mesh the sum of all the iR drops must equal the sum of all the electromotive forces taken with due regard to sign.* This simply says that if we start at some initial point in a mesh, if we take all the iR drops in that mesh in going around the mesh, they must equal the sum of the electromotive forces in the mesh, for the circuit only contains iR drops and sources of potential and if the sum of the iR drops and E.M.F.'s going around the mesh were not equal, we would come back to the initial point with a potential which is different from that with which we started, and a point cannot be at two different potentials at the same time. Symbolically, for the mesh $AEBC$ calling the currents in the branches in the figure above, AB, BC, CA, i_{AB}, i_{BC}, i_{CA}, and resistances R_{AB}, R_{BC}, R_{CA}, and calling E the E.M.F. at the cell and R_i its internal

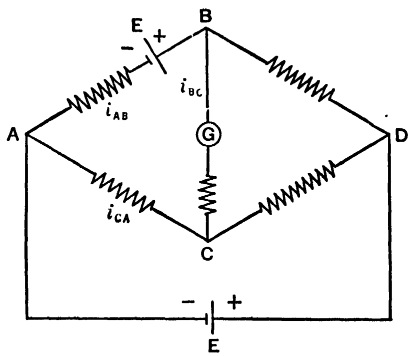


FIG. 35.—Illustration of Meshes in a Complex Circuit.

resistance, we have $i_{AB}R_{AB} + i_{BC}R_{BC} + i_{CA}R_{CA} + i_{AB}R_i = E$. For the general case, Kirchhoff's second law would be expressed by writing $\Sigma iR = \Sigma$E.M.F.*

These two laws as we shall see give us a chance to write a number of equations. The first set of equations will be the equations for the equality of current flow towards and away from every junction point in the circuit. This series of equations expresses Kirchhoff's first law. A second series of equations will be obtained by writing the second Kirchhoff's law for each of the possible imaginable meshes in the circuit. From the study of the number of meshes and points in any circuit it will be seen that one will have enough algebraic equations resulting from these two investigations to enable one to solve for all the unknown currents in the system.† An example of how one proceeds we can obtain from a study of the Wheatstone bridge. This instrument is one which is almost universally used for comparing resistances.

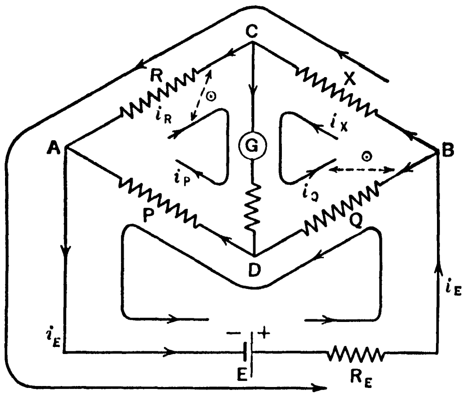


FIG. 36.—Application of Kirchhoff's Laws to Wheatstone's Bridge.

In the circuit outlined in Fig. 36, X is an unknown resistance whose value R_X is to be compared with that of R, or R_R, which is a standard adjustable resistance. The resistances P and Q, of value R_P and R_Q, are resistances whose values have been adjusted by trial and error so that they are in the approximate ratio of the resistances R and X. G is a galvanometer of resistance R_G and E is the battery sending a current through the system. Call i_R, i_X, i_P, i_Q, i_G, and i_E,

* If we include the internal resistances in the iR drops, the equation is that written $\Sigma iR = \Sigma$E.M.F. For simplicity in working problems we shall in subsequent treatment regard the P.D. at the cell terminals and the external iR drops only. Although if the E.M.F.'s and R_i's values are given the treatment is identical. The use of the E.M.F.'s and R_i's merely adds more resistance terms and increases the chance for numerical blunders.

† It is unnecessary to write a set of equations for every possible mesh. If this be done a number of equations will be found to be merely the addition of two or more equations already given. By careful study enough independent equations can be chosen to enable one to solve for each one of the unknowns. This is not difficult to accomplish for it is only necessary to write enough of these equations to contain each iR or E.M.F. of the entire circuit at least once. The same holds true for each of the i's of the first law.

the currents in the branches of the circuit having the resistances R, X, P, Q and in the galvanometer G and through the battery resistance R_E.

Before applying the laws it is necessary to *assume* the direction of the flow of currents in this system. It is very possible that the assumptions as to the direction of flow under a given set of conditions *may be wrong.* The error, however, on proper solution of the equations, will be readily *found out* in that the current whose direction was wrongly assumed will come out *negative.* This is an automatic property of the equations and all that one needs in fixing the sign, that is the convention for the flow of the current, is that a *definite system be chosen and that that be adhered to in all the calculations.* If the current comes out negative, then in discussing the circuit after solution the correction must be used. The small arrows in the wires of the diagram indicate the arbitrary directions of flow assumed. It may be added that the direction of flow through the branches BDA and BCA are fixed by the single battery in the circuit. The assumption lies only in the direction of the current flowing through G, and an error in this assumption can be found only when numerical values are substituted and the equations solved. *With the choice of the flow of current* one may at once write the equations of current flow for the 4 points A, B, C and D from Kirchhoff's first law.

At	A	$i_E = i_R + i_P$
	B	$i_E = i_X + i_Q$
	C	$i_X = i_G + i_R$
	D	$i_P = i_G + i_Q$

Before writing the equations based on the second law it is *essential that we pick the meshes and indicate the direction of the flow of current in the mesh.* This is required in order that we know definitely the sign to assign to various terms in the equation. The meshes will be indicated by the circular arrows in the diagram, and the direction of current flow is taken as positive in the sense of the arrow. The mesh equations may now be written down. In the mesh

$EBCA$	$\text{P.D.}_E = R_E i_E + R_X i_X + R_R i_R$
$EBDA$	$\text{P.D.}_E = R_E i_E + R_Q i_Q + R_P i_P$
BCD	$0 = R_X i_X + R_G i_G - R_Q i_Q$
CDA	$0 = R_G i_G + R_P i_P - R_R i_R$

In the first two meshes the *directions of the assumed currents* also *follow the circular arrows drawn* indicating the sense of the positive flow of current. Consequently, the law can be written that the potential of the cell E is equal to the sum of the positive iR drops in the circuit. In the two triangular meshes, however, the *arrow indicating the positive sense of current flow* assumed, *points in a contrary direction to the assumed current flow in the case of the wires BD and AC.* Hence the sense of the iR drops corresponding to R_Qi_Q and R_Ri_R must be negative. The fact that there is no cell in these two meshes makes the sum of the iR drops equal to 0, as the P.D. in each mesh is 0 in the absence of a cell. With these eight equations it is easy to eliminate the six unknown variables and so solve for the value of any one of them. In fact all that is needed is a set of six *independent* simultaneous equations containing the six unknowns which will enable one to solve the equations. The remaining two equations deduced can serve as checks in the computation. The solution of *complicated* simultaneous equations involving *many* variables can be most easily accomplished by means of determinants. For the use of determinants one is referred to any standard college algebra. Often determinants are not necessary in the solution of problems and the equations can be solved by inspection.

35. THE USE OF THE WHEATSTONE BRIDGE

In the ordinary use of the Wheatstone bridge the variable resistance R_R is changed in value until the current through the galvanometer i_G is 0. In this case, the solution of the equations above is much simplified and one at once finds the law of the Wheatstone bridge as commonly used. Set $i_G = 0$, then the equations for meshes CBD and CDA become:

$$R_Xi_X = R_Qi_Q$$

and

$$R_Pi_P = R_Ri_R.$$

Dividing R_Qi_Q by R_Pi_P, one has this relation:

$$\frac{R_Qi_Q}{R_Pi_P} = \frac{R_Xi_X}{R_Ri_R}.$$

Since i_G is 0, the equations for the points C and D give us, $i_X = i_R$, and $i_P = i_Q$. Hence

$$\frac{R_Q}{R_P} = \frac{R_X}{R_R}.$$

Thus, when a balance is obtained on the bridge the resistance R_X of

the resistance coil X is equal to the resistance R_R of the known coil R multiplied by the ratio of the resistances of the coils P and Q.

36. NOTE ON THE SOLUTION OF KIRCHHOFF'S LAW EQUATIONS

The procedure outlined below is not urged on the student but experience has shown the writer that the care and precautions used below lead to a successful solution, while if they are omitted this is not the case. Students often balk at the trouble of drawing in current and potential fall arrows in simple circuits. The chance for numerical blunders in the solution of the equations is quite enough of a hazard without the danger and waste of time in solving erroneous equations. The plan outlined below if followed out with care is as nearly " fool proof " as any plan can be and if adhered to will save many hours.

In general for a practical solution of the currents in a circuit by Kirchhoff's laws of divided circuits we may proceed as follows:

(1) Make a neat, clear diagram of the circuit to be analyzed designating all the resistances, E.M.F.'s or P.D.'s by appropriate symbols that cannot be misread. The diagram should be large enough so that the operations to be carried out below can be clearly indicated. The diagram should have all junction points of wires clearly labeled.

(2) Inspect the diagram and from the values of the E.M.F's or P.D's and resistances make a consistent *guess* as to how the currents flow along the wires. Indicate your *chosen sense of flow by arrows placed on the wires*, and check to see that no single wire has two oppositely pointed arrows on a length between the two junction points of the circuit defining that wire.

(3) Then assign a *consistent fall of potential to each of the meshes in the diagram* so that there is not a single wire that has not an indication of the sense of potential fall assumed. To do this start with the positive pole of the highest E.M.F. or P.D. represented and proceed around the circuit. Where there are two E.M.F.'s or P.D.'s opposed the sense of fall of potential can be chosen starting *from* the positive pole of the highest E.M.F. or P.D. around the mesh. These *potential falls* in the meshes may be indicated by circular arrows, possibly in some other color than the lines in the drawing. The circular arrows may be drawn inside or outside of the mesh in question wherever most convenient and where they do not overlap other arrows.

(4) Next, taking the junction points of the wires in order, write down the equations of current flow for each one separately, in either one of the two following ways: All currents *flowing toward* the point can be set on the *left hand side* of an equal sign having on its right hand

side all currents flowing away from the point, or the equations may also be written with all currents flowing *toward* the junction point having a *positive sign* and all those flowing *away* from the point having a *negative sign,* the sum of all these currents being equated to zero. This method is the one indicated by the equation $\Sigma i = 0$.

One should next check over the equations, and of those that are identical or are the result of the addition of two or more of the equations already written, strike out such as to obtain a complete set of *independent* current equations.

(5) Next proceed to write the equations for the various meshes. In writing the equations label the mesh chosen and write its equation putting on the left hand side all E.M.F.'s or P.D.'s and on the right hand side of the equality sign all the iR drops. In doing this be *careful to watch signs.* Wherever an E.M.F. or P.D. in the circuit is *contrary* in sense to the *circular arrow chosen to indicate the fall of potential place a negative sign in front of it.* Wherever the current runs in a sense *counter to the fall of potential indicated by the circular arrow the iR drop in the branch having this current should be preceded by a minus sign.* It may be advantageous to inspect the circuit before writing the equations and indicate by a distinguishing mark all places where contrary signs are found.

After writing all mesh equations inspect the equations and see that some of these are not simple combinations of other meshes. *It is only necessary that each wire or E.M.F. be counted in one mesh to insure enough equations for a solution.*

(6) Having now the current and mesh equations complete count the total equations involved. There should be as many independent equations as there are unknown currents to be evaluated. The inclusion of additional equations which are combinations of those present will only complicate the solution and lead, if carelessly used, to identities that are annoying. If the numerical values of the E.M.F.'s or P.D.'s and resistances have not been introduced and the equations have been set up symbolically only, then introduce the values of the resistances. The equations are now ready for solution either by determinants or by simple algebraic manipulation.

(7) If the equations are to be solved directly a great deal of time in aimless algebraic computation may be saved by a careful inspection of the equations leading to an outline of the proposed solution. Usually only one of the unknown currents is required. The objective is to solve for this one current. It is then necessary merely to regard the equations containing this particular unknown and plan a series of combinations of the equations so as to systematically eliminate

all other unknown currents, finally obtaining an equation containing but the one unknown. While with equations containing more than five unknowns this seems a formidable task, the difficulties are not great if headwork and care are used in planning the elimination.

(8) When the equations are solved some of the currents may appear with a ***negative*** sign. This means that ***the direction of the current flow is opposite to that assumed.*** The equations can be checked by taking any one of the combination equations for some of the meshes and substituting the values obtained.

A practical example of the application of Kirchhoff's laws may be made to the circuit shown in Fig. 37. Let E_1 and E_2 be batteries producing potentials of 4 and 2 volts respectively, while the resistances R_1, R_2, and R_3 have the values 2, 2, and 4 ohms respectively. It is desired to know the currents i_1, i_2, and i_3 in the branches of the circuit corresponding to the resistances R_1, R_2, and R_3. To solve this problem we proceed as follows. Inspection of the circuit shows that E_1 has the highest potential and the presumption is that the current from E_1 will flow in the direction of the small arrows indicated. At the point A the current will divide, one part going into R_3 and the other part going into R_2. Were E_2 large and R_2 small it is possible that the current from E_2 would flow through R_3 in the opposite sense from that indicated. It is therefore not certain exactly how the current through R_3 is going to flow. We simply make the arbitrary *assumption* that it flows from A to B, basing our judgment on the fact that E_1 is greater than E_2. If it should turn out that the current flows in the opposite sense we will obtain a negative value for i_3 in the result. Having made this assignment of current flow we can at once proceed to write the equations for the currents at the junction points A and B.

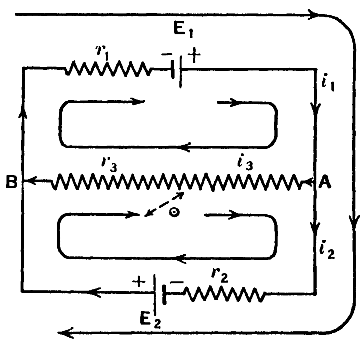


FIG. 37.—Application of Kirchhoff's Laws.

At $\qquad A \quad i_1 = i_2 + i_3$

$$B \quad i_2 + i_3 = i_1.$$

With this simple circuit it happens that the two current equations at the symmetrical junction points A and B are the same. Hence only one of these equations is necessary since either one contains all

of the currents once. This is not usually the case in more complicated circuits.

Again regard the meshes involved. We have the mesh containing R_1, E_1, and R_3. Since the current flows from E_1 and in the direction assumed in R_3 we can assume that the circled arrow represents the direction of positive current flow. In the mesh $R_3R_2E_2$ the direction of positive current flow may be assumed to be in the direction of the current arrow again. Finally, in the mesh $R_1E_1R_2E_2$ the sense of positive current flow is indicated by the encircling arrow. Having assumed a reference system of current flow we may write the second set of equations. In the mesh $E_1R_1R_3$, $i_1R_1 + i_3R_3 = E_1$. In the mesh $E_2R_2R_3$, $i_2R_2 - i_3R_3 = E_2$. In the mesh $E_1E_2R_1R_2$, $i_1R_1 + i_2R_2 = E_1 + E_2$. Again it is to be noted that only two of these equations are independent. It now remains to put in the numerical values for the resistances and solve the equations.

In the $E_1R_1R_3$ mesh substitution of the values gives $2i_1 + 4i_3 = 4$.
In the $E_2R_2R_3$ mesh substitution of the values gives $2i_2 - 4i_3 = 2$.
In the $E_1E_2R_1R_2$ mesh, $2i_1 + 2i_2 = 6$.

Therefore

$$i_3 = 1 - \tfrac{1}{2}i_1.$$

Also

$$i_2 = 3 - i_1.$$

From the fact that $i_2 + i_3 = i_1$, we have $i_1 = 1 - \frac{1}{2}i_1 + 3 - i_1$, hence

$$i_1 = \tfrac{8}{5}.$$

Placing this in the equation for i_2, we have $i_2 = \frac{7}{5}$, and $i_3 = \frac{1}{5}$.

In this calculation we have made use of equations for mesh $E_1E_2R_1R_2$, and the $E_1R_1R_3$ mesh. The correctness of our calculation can be checked against the equation for the $E_2R_2R_3$ mesh. This says that

$$i_2R_2 - i_3R_3 = E_2.$$

Putting in the values for R and E the equation becomes $2i_2 - 4i_3 = 2$. The values for i_2 and i_3 are $\frac{7}{5}$ and $\frac{1}{5}$, respectively. Placing these in the equation we see that $\frac{14}{5} - \frac{4}{5} = 2$. Our calculation has therefore been checked. If the resistances were in ohms the currents are given in amperes.

CHAPTER IX

ELECTRICAL MEASURING INSTRUMENTS

37. MOVING COIL OR D'ARSONVAL GALVANOMETER

OF all the electrical measuring instruments probably the most important, and the one whose theory applies to most of the current and potential measuring instruments now in use, is the moving coil or d'Arsonval galvanometer. In Chapter V, it was shown that from Ampère's rule we could define unit current in two ways. The second definition was that *unit current is that current which flowing in unit length of conductor perpendicular to a uniform magnetic field of unit strength experiences a mechanical force of 1 dyne.* This definition of unit current may be expressed in the statement $f = i_a Hl$, where l is the length of conductor, H is the strength of the uniform field, i_a is the current and f the mechanical force in dynes. Assume we have a coil suspended in the uniform magnetic field shown in the Fig. 38. The coil may be of rectangular cross section and be so placed that the wires are parallel to the pole pieces of the magnet causing the field. Assume that there are n turns of wire in the coil and that the length of the coil is l cm and its breadth is d cm. The force for the n wires which run up one side of the coil when a current of i_a absolute units runs in the coil is given by

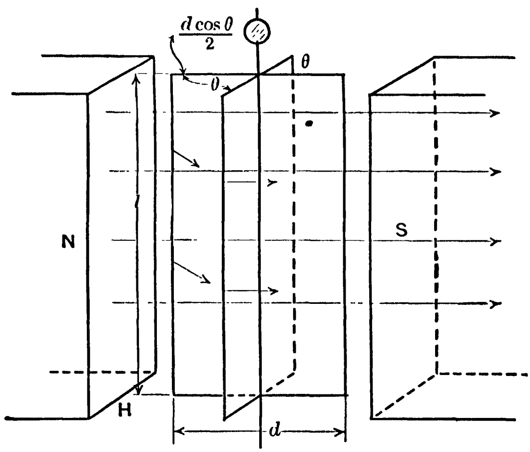


FIG. 38.—Schematic Diagram of the Moving Coil Galvanometer.

$$f = ni_a Hl.$$

A similar force will be experienced by the other n vertical wires. The force moment due to the n wires is $G_1 = ni_aHl\left(\frac{d}{2}\cos\theta\right)$ if the plane of the coil has been rotated through an angle θ with the magnetic field. This is obvious since the lever arm of the wire when the coil lies parallel to the field is $\frac{d}{2}$, and if it is displaced through an angle θ, the lever arm of the force acting will be the projection of $\frac{d}{2}$ on the direction of the field. Since there are two wires, the total force moment will be twice that given above, namely: $2G_1 = G = ni_aHld\cos\theta = Ai_aH\cos\theta$. The quantity $A = nld$ is a constant of the coil, and is merely the total area enclosed by the n turns. The wires of length d in the coil are parallel to the field and therefore have no force exerted on them as regards G. If the coil had been of any other shape, the result would have been the same, as any area A can be broken up into rectangular elements of height dl and length d which now varies along the coil. The force on each element will again depend on $ndld$, and the total force will depend on $\int nddl$ = the area of the coil.

Now the coil is suspended by a wire and the force moment on the coil due to the field acts to turn the coil at right angles to the field against the elastic forces of the suspension. For a deflection θ of the galvanometer, G is equal to $T_0\theta$. (See Chapter IV.) Thus $T_0\theta = Ai_aH\cos\theta$. If small deflections are measured then $\cos\theta$ approaches unity, so that for deflections less than 10° we can use the approximate equation.

$$i_a = \frac{T_0\theta}{AH}.$$

The use of small deflections is simplified by the fact that a mirror hung on the coil will give an appreciable deflection of a spot of light for a small angular deflection, if the distance of the source of light and the spot from the mirror is sufficiently great.

38. THE GALVANOMETER CONSTANT

The equation for the current just deduced is the equation underlying all moving coil galvanometers. The ratio $\frac{T_0}{AH}$ is a constant of the apparatus as long as (1) the area of the coil remains constant, (2) the magnetic field due to the permanent magnets remains con-

stant, and (3) the constant T_0 of the suspension is not changed. It is therefore called the galvanometer constant K', and we can write $i_a = K'\theta$, where θ is in radians. The response of a galvanometer to a current is determined by having a large θ for a small i_a, that is, in having a large deflection for a small current. Thus, if θ is large when i_a is small, K' is a small fraction, so that for large response to a weak current $i_a/\theta = \dfrac{T_0}{AH} = K'$, and the constant factor is small. It is obvious that the K' will be smaller the greater A, the greater H, and the smaller T_0.

If one make T_0 too small relative to the inertia of the moving coil, the period of the galvanometer becomes very long and it is difficult to work with. Increasing A beyond certain limits also increases the period, and too strong a field is difficult to maintain constant. Hence there are limitations to the sensitivity which may be obtained. It is generally best to use many turns of fine wire on a coil which has a great l and a small d. Therefore the moment of inertia is decreased and a finer suspension may be used. The constant K' is rated in terms of the current i_a which is necessary to give one radian of deflection of the mirror. That is, the galvanometer constant

$$K' = \frac{K}{10} = \frac{i}{10\theta} = \frac{T_0}{AH},$$

where *i is the current in amperes* required to give θ *radians* deflection.

It is usually simpler to speak of what is known as the figure of merit of the galvanometer. This is the current in amperes necessary to cause a deflection of 1 mm on a scale distant 1 meter. This figure of merit k is related to the galvanometer constant K from geometrical considerations by the expression

$$k = \frac{K}{2000}.$$

This follows since the spot of light is deflected through θ when the mirror turns through $\theta/2$ radians, and since a deflection of 1 mm at 1000 mm distance gives tan $\theta/2$ which approaches $\theta = \frac{1}{2000}$. Therefore the current to cause a deflection of 1 mm on a scale at 1 m is $\frac{1}{2000}$ the current to cause a deflection of the mirror through 1 radian. Whence one has at once that the figure of merit

$$k = \frac{1}{2000}\left(\frac{i}{\theta}\right) = \frac{K}{2000}.$$

The term sensitivity is also often used. It is expressed by a unit

called the megohm.* The megohm is merely the resistance in millions of ohms which placed in series with the galvanometer will, when 1 volt potential difference is placed across the resistance of the galvanometer, cause a deflection in the galvanometer of 1 millimeter at a meter's distance. The word megohm refers to resistances of 10^6 or millions of ohms as the unit. This unit is the reciprocal of k, divided by a million. The sensitivity in megohms is therefore equal to

$$\frac{1}{k10^6} = \frac{2000}{K}/10^6 = \frac{1}{500K}.$$

Thus if $K = 0.000004$ when the current i is in amperes, the figure of merit $k = 0.000000002$.

Hence, $\frac{1}{k} = 0.5 \times 10^9 = 500 \times 10^6$, so that the sensitivity therefore is 500 megohms.

39. TYPES OF GALVANOMETERS, AMMETERS AND VOLTMETERS

For most current measuring work today the galvanometers used are of the moving coil type. Such galvanometers have been developed to a high degree of sensitivity and precision, and currents of the order of magnitude of 10^{-11} amperes may be measured with them. For still weaker currents the galvanometers used are generally of the suspended needle type. In these the coil carrying the current generates the magnetic field and consists of many thousands of turns of fine wire. The magnets are as a rule small magnetized strips of steel. For the most sensitive ones these galvanometer magnets are arranged in pairs of two, the north pole of one magnet of the pair being opposite to the south pole of the other. These give very weak magnetic fields which are not subject to the action of the earth's field to any extent. Such galvanometers are known as *astatic.* They have been developed to a high degree of sensitivity for use with delicate thermocouples to measure the intensity of the heat radiation from the stars. They are also used in measuring the intensity of light in the various spectral lines. Such galvanometers must be protected from the earth's magnetic field, and this is accomplished by placing a series of soft iron shields around them. In one used by Professor Millikan, which has been designed by Dr. Coblentz at the Bureau of Standards, seven such shields of soft iron were required to cut out

* Other ways of expressing sensitivity are also used, but omitted for the sake of clarity. (See Laws, Electrical Measurements.)

the earth's field. The limit of sensitivity in galvanometric measurements is about 10^{-12} amperes.

The galvanometer, as can be seen above, is used for measuring currents, and the deflection given is proportional to the current. Since, however, the current through the galvanometer is proportional to the voltage across the galvanometer it is possible to calibrate it to read voltage. *In order that a galvanometer should read voltage all that is needed is that the current carried by the galvanometer should be so small compared to the current carried by the portion of the circuit across which the potential difference is to be measured that connecting the galvanometer shall not materially change the value of the current in the rest of the circuit,* and thus alter the potential.

Consider a wire of resistance R, through which a current i flows. It is desired to know the potential difference existing across this wire. If a galvanometer of resistance R_g, which is very large compared to R, be placed across R, the current through the galvanometer will be so small as to leave the potential drop across R unchanged. The current through the galvanometer then will be practically proportional to the potential which existed across R before the galvanometer was connected in. If the deflection of the galvanometer in terms of the potential causing that deflection can be determined by calibration the galvanometer will read voltage.

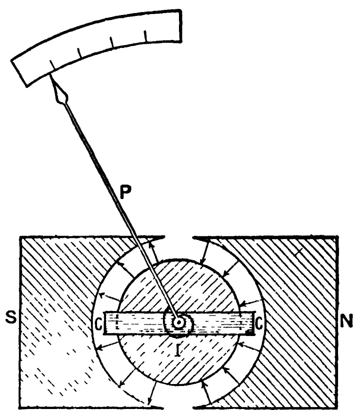


FIG. 39.—Schematic Diagram of a Direct Current Voltmeter or Ammeter.

Thus a galvanometer may be used in two ways. If it has a *low resistance* so that *its series resistance* does not materially increase the resistance of the circuit into which it is introduced, it will read the *current* in the circuit, that is, it acts as a galvanometer or an *ammeter.* If it has a very *high resistance* compared to the *resistance of a portion of the circuit across which it is introduced,* it will then read *the potential difference* of that part of the circuit across which it is connected.

For commercial purposes for measuring potential and current the galvanometers are modified in the following fashion. The coil is mounted on a vertical axis on jeweled bearings. The suspension is replaced by a small spiral spring. The coil has the same shape as before but the pole pieces are very ingeniously arranged as shown

in Fig. 39, so that they give a uniform field perpendicular to the coil through a range of deflection of nearly 90°. The form of the pole pieces to accomplish this is very simple and is shown in cross section in the diagram. The poles of the large permanent magnet are labeled N and S. They have a cylindrical hole drilled perpendicular to the plane of the paper. In the center of this hole is a soft iron lug I concentric with the axis of the cylinder. The small rectangles labeled C and C constitute cross sections of the vertical portions of the coil. The arrow P is a pointer which moves over a uniformly divided scale reading volts or amperes directly. It is obvious that the soft iron lug causes a radial field to pass through it over a limited range of angles. Thus, as the coil rotates about its axis it stays in a nearly uniform field, so that the deflection is proportional to the current, for the restoring force of the spring is greater the greater the deflection. The instruments described above are called respectively ammeters or voltmeters, depending on their resistance, which determines their use as above stated.

It must be pointed out that such instruments, while they register current and voltage, cannot be used as *absolute* measures of current and voltage. Every voltmeter or ammeter put on the market *must be calibrated* in terms of the standard of current or potential.

In the electromagnetic system of units the standard of potentials was defined in terms of the heating effect of the current; that is, fixed points on the scale of the voltmeter can be determined by observing the heating effect produced in a wire across terminals of which the voltmeter is connected. Again the ammeter must be calibrated to read amperes and the calibration will be effected by means of a standard based ultimately on the measurements of the tangent galvanometer. As will be remembered, this again depends for its absolute value on the magnetometric measurements of the earth's field.

It is to be noted that for galvanometers and voltmeters a given calibration holds good only as long as no change in the instrument has taken place. Although the field magnets on most galvanometers are so arranged as to maintain their magnetism for long periods of time the magnetism of such galvanometers must obviously change, in general decreasing in the course of years. No ammeter or voltmeter should be used without having checked its scale against a standard within a year.

It is possible to arrange voltmeters or ammeters in which the current flows not only through the armature coil but also around coils which are wound on the soft iron which gives the field, as shown in

Fig. 40. In this case, both H, the field, and the torque, G, are caused by the same current. The deflection θ is therefore proportional to i^2. Such instruments have the advantage that they are *independent of a variation in the magnetic field.* They have, however, a *scale in which the deflection varies as the square of the current.* Since the currents in the field coils and the armature or suspended coil are always in the same direction relative to each other the torque will always be in the same direction even if the current through the line changes in sense. Inasmuch as industrial development has led to the universal use of alternating currents which vary in direction with time, practical needs have demanded the widespread use of this type of ammeter and voltmeter. The interpretation of the meaning of the reading of such instruments must be reserved for Chapter XXIV, where alternating currents are discussed. Beside the use of such instruments for alternating currents, the necessity of non-varying standards makes these instruments of value, for the other type of instrument has permanent magnets whose strength may vary with time.

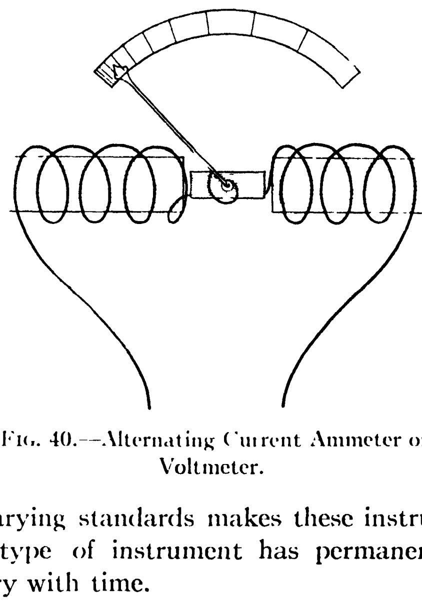


FIG. 40.—Alternating Current Ammeter or Voltmeter.

40. SHUNTS FOR AMMETERS, SERIES RESISTANCES FOR VOLTMETERS

The non-uniform scale divisions in ammeters and voltmeters which are used for standards, and which are based on the principle discussed at the end of the last section, where the field is excited by the current which flows through the armature, limit the range for accurate measurement on any given instrument. Such instruments are often used as standards, and to increase the range of instruments of this type we make use of a device known as the shunt, in the case of ammeters. In the case of voltmeters, a device equivalent to the shunt is used. It is worth while briefly to discuss the shunt, the theory of which follows directly from Ohm's law. Assume an ammeter, represented by A in Fig. 41, which has a resistance R_A. Now

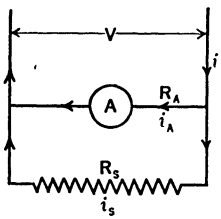


FIG. 41.—Diagram of an Ammeter Shunt.

place a resistance R_s in parallel with it. Then, as the potential difference is V, the current i_A flowing through the ammeter, is given by

$$i_A = \frac{V}{R_A},$$

and the current i_s through the shunt is given by

$$i_s = \frac{V}{R_s}.$$

Finally, for the two combined, $\frac{V}{R} = i$, where i is the current through both ammeter and resistance, and R is the resistance of the combination of R_A and R_s. From the law of resistances in parallel,

$$\frac{1}{R} = \frac{1}{R_A} + \frac{1}{R_s}.$$

Therefore,

$$R = \frac{R_A R_s}{R_A + R_s}.$$

Hence,

$$\frac{V}{\dfrac{R_A R_s}{R_A + R_s}} = i.$$

Thus

$$\frac{i_A}{i} = \frac{V}{R_A} \Bigg/ \frac{V}{\dfrac{R_A R_s}{R_A + R_s}},$$

or

$$\frac{i_A}{i} = \frac{R_s}{R_A + R_s}.$$

This states that the current through the ammeter is to the total current through the circuit as the resistance of the shunt is to the sum of the resistances of ammeter and shunt. Thus, if one reads the current given by the ammeter and multiplies it by the ratio of the sum of the resistances of ammeter and shunt to the resistance of the shunt, one has i the current flowing through the whole circuit. If then one has an ammeter which reads over a range of 1 ampere and one wishes to make it read over a range X times as great, all one need do is to make a shunt of resistance such that the resistance R_A of the ammeter

plus the resistance R_s of the shunt is equal to X times the resistance R_s of the shunt. Under these circumstances the ratio

$$\frac{R_A + R_s}{R_s} = X$$

$$i = Xi_A.$$

This means merely that if one reads the current, through the ammeter as given by it, and multiply this reading by the factor X, one at once has the current i flowing through the circuit.

To increase the range of a voltmeter we make use of a principle which follows simply from Ohm's law. Assume a voltmeter with an internal resistance R_V, and place in series with it a resistance R_s, as shown in Fig. 42. Call R the combined resistance of R_V and R_s; that is

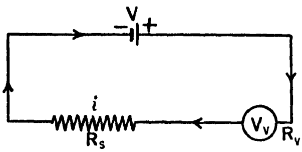


FIG. 42.—Series Resistance in a Voltmeter Circuit to Extend its Range.

$$R = R_V + R_s.$$

The current i, through the circuit, if a potential V is applied across the two resistances in series, is given by

$$i = \frac{V}{R} = \frac{V}{R_s + R_V}.$$

The potential difference V_V across voltmeter gives a current i which is the same as the current flowing through the system as a whole. Thus

$$\frac{V_V}{R_V} = i$$

From the above equation one has

$$\frac{V_V}{R_V} = i = \frac{V}{R_s + R_V},$$

so that

$$\frac{V_V}{V} = \frac{R_V}{R_s + R_V}.$$

This says that the potential difference across the voltmeter, which is the potential difference registered by it, is to the total potential difference across the voltmeter and the resistance R_s, which we denoted by V, as the resistance of the voltmeter is to the sum of the resistances R_s and R_V. If we wish to use a voltmeter which reads up to 150 volts

to read 150 (X) volts, all we need do is to make the sum of the resistances R_s and R_V equal to X times the resistance R_V. Then

$$\frac{R_V}{R_s + R_V} = \frac{1}{X},$$

and

$$\frac{V_V}{V} = \frac{1}{X},$$

or

$$V = X \cdot V_V.$$

This states that if we have a potential across the voltmeter of a given amount, the potential across the combined voltmeter and the resistance R_s is X times as great. For example, if we wish to read a voltage of 1500 volts on the voltmeter above, we place in series with the voltmeter a known resistance whose ratio to the resistance of the voltmeter is given by

$$\frac{R_V}{R_s + R_V} = \frac{1}{10}.$$

Then if the voltmeter reads 90 volts the potential will be 900 volts.

41. THE WATTMETER

The principle which underlies the voltmeter and the ammeter may be made use of to measure power directly. In Chapter VI we defined the practical unit of power, the watt, as being obtained by multiplying the amperes flowing through the circuit by the potential difference in volts which is maintained across the circuit. As was stated before, a moving coil galvanometer with a high resistance acts like a voltmeter. If the magnetic field H of the galvanometer be produced by a coil of wire through which the main current in the circuit is flowing, the field H will then be proportional to the current i. The force on the moving coil will be the product of the field H and the current flowing through the high resistance coil. But this current measures the potential, V, across the coil, and H depends on i, the current through the low resistance system. Thus the torque produced will be proportional to the product of the current and the potential across the circuit, iV; that is, the deflection will be proportional to the power in the circuit. Such an instrument is known as the wattmeter.

A schematic diagram of the instrument and its relation to the circuit is shown in Fig. 43. In the diagram the magnetic bars of soft iron N and S are magnetized by the coils F. The current which flows

through these is the main current flowing through the line. M is a motor whose power consumption is being measured. It is connected so that as will be seen M is in series with the coils. The small armature coil A of the wattmeter has a very high resistance. The current through it therefore does not affect the current through the motor or the rest of the line materially. Its terminals connect it to the two terminals of the motor. The current through it is therefore proportional to the potential across the motor. The deflection of A depends on the product of the potential across A and the current i through the coils F producing the fields. The instrument reads watts only if it is arbitrarily calibrated by comparison with standard instruments. Where the currents vary periodically with time, and the potential difference and current in the motor are not in phase with each other (see Chapter XXV), the wattmeter described above gives the only means of measuring the true power consumption, for it automatically gives a deflection proportional to the average of the product of current and potential in the phase relations existing in the circuit.

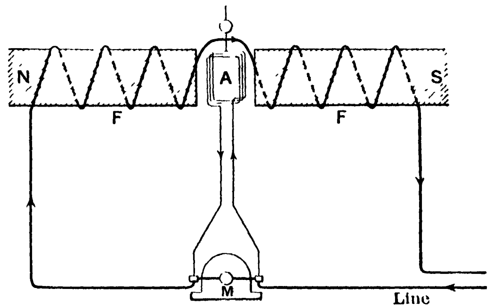


FIG. 43.—Schematic Diagram of a Wattmeter.

CHAPTER X

MEASUREMENT OF CURRENT, POTENTIAL DIFFERENCE, AND RESISTANCE

42. MEASUREMENT OF CURRENT

IN the previous chapter, the measurements of currents and potential differences by means of the magnetic effect of a current were given. In that chapter the fundamental principle of such a measuring instrument as the moving coil galvanometer was discussed. It is now necessary to discuss other means of measuring currents and potentials. Currents may be measured by the following devices:

(1) Tangent galvanometer.
(2) Any calibrated galvanometer of the moving coil or magnet type.
(3) The hot wire ammeter.
(4) The potentiometer, in combination with a standard resistance.
(5) By static methods for small currents.
(6) By electro-deposition of metals applying Faraday's laws of electrolysis.

(1) The measurement of current by the tangent galvanometer has been discussed in Chapter V. In this method, the tangent galvanometer, the constant of which is known, is placed with the plane of its coil parallel to the earth's magnetic field. The current is passed through the coil, and the angle of deflection of the needle is noted. The formula there given

$$i = \frac{10rH}{2\pi n} \tan \theta,$$

will give the current i in amperes from the observed angle θ and the constants r, H and n. The instrument is not particularly accurate inasmuch as the field H cannot be accurately determined. It is also awkward to use on account of the continual variation of the earth's field. It is, however, the standard of comparison for all measuring

instruments which give i_a on the basis of the absolute electromagnetic system. Today however currents are more accurately defined by means of the current balance. This measures the force in grams between two coils having the same current through them, the form of the coils being such that the force may be computed in terms of the electromagnetic definition of current in the absolute system.

(2) The measurement of currents by galvanometers has been discussed adequately in Chapter IX.

(3) Hot wire ammeters are ammeters which depend on the heating effect produced by the current. The heating produced by the current is proportional to the square of the current. In some cheap instruments the heating produced by the current merely changes the tension of the wire through which it flows, which acts as a spring activating a needle by an appropriate lever system. Another form of the hot wire ammeter which is used frequently in radio instruments is one in which the heating of the wire by the current warms up a thermal junction fastened to the wire. This thermal junction then gives an E.M.F. which can be amplified and detected by an appropriate device. Instruments of this type are mainly used for measuring high-frequency alternating currents, where iron cannot be used and the currents are so small that the mechanical forces obtained are difficult to observe.

(4) The potentiometer may be used for measuring the small currents across an accurately known resistance, inasmuch as the potentiometer gives the potential across the resistance, and division of the potential by the resistance at once gives the current flow. In this case measurement is effected accurately without altering the current, as the potentiometer embodies a no-flow method.

(5) When we study static electricity, we shall learn that the quantity of electricity on a system can be defined by the product of the capacity of the system and the potential to which it is raised. By measuring the change of potential of a system of known capacity in a given time the gain or loss of quantity per unit time can be easily determined. By this means currents as small as 10^{-17} amperes have been measured using the Hoffman electrometer.

(6) When a current flows through a solution of some electrolyte there is a liberation either of metallic ions or of gases at the electrodes. Faraday's laws of electrolysis definitely specify the weight of material so liberated as related to the quantity of electricity which has passed and certain chemical characteristics of the substances liberated. If the weight of substance liberated in a given time is known by means of these laws, which will be elaborated in Chapter XI, the current can be easily calculated. The legal unit of current is based on the deter-

mination of the rate of deposition of silver under controlled conditions utilizing this principle.

43. POTENTIAL MEASUREMENT

E.M.F., or better potential difference, is measured by the following instruments:

(1) The voltmeter or high-resistance calibrated galvanometer.
(2) By the potentiometer.
(3) By the electrostatic voltmeter.
(4) By heating produced for a known current.

(1) The measurement of potential by the first method has already been discussed in Chapter IX.

(2) The theory of the potentiometer which was developed in Chapter VII enables one to determine the potential by comparison with a known potential.

For comparing electromotive forces we make use of the potentiometer shown in Fig. 44. At the point V in the circuit any battery of *constant* potential larger than the batteries whose potential it is required to compare is used. G is a galvanometer inserted in series with one of the batteries to be compared. These batteries are labeled V_s for a standard cell or V_x unknown. The arrow indicates a variable resistance contact which divides the resistance R_1R_2 into two parts, the ratios of the resistances R_1 and R_2 of which can be accurately determined. The law of fall of potential down a resistance wire states that if the potential due to the two cells at P is the same, the potential difference of the standard must be to the potential V as R_2 the resistance across the standard is to $R_1 + R_2$. If G shows no deflection, the condition of equal potential is fulfilled. If the ratio of the resistances is not properly adjusted the current will flow through the galvanometer G in one sense or the other. By careful adjustment of resistance until no current can be detected an accurate measurement of the potential in terms of the standard cell and the resistances can be obtained.

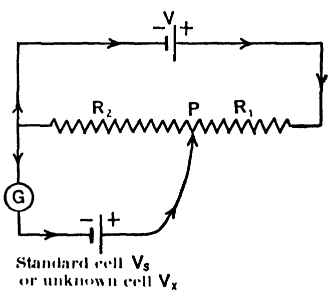


FIG. 44.—Comparison of E.M.F.s by Use of the Potentiometer.

$$\frac{V_s}{V} = \frac{R_2}{R_1 + R_2}.$$

In practice the standard cell is placed in position and the resistance ratio of R_1 to R_2 is changed until no current flows through the galvanometer. The standard cell V_s is then removed and the unknown cell V_x is put in its place. Adjustment is again made until the balance is obtained with no current in the galvanometer. In this case, the ratio of the resistances will be different and will be represented by resistance R_3 and R_4, where R_3 replaces the value of R_1, and R_4 replaces the value of R_2 obtained with the standard cell. If the potential difference produced by the unknown cell be designated by V_x then when no current flows in the galvanometer

$$\frac{V_x}{V} = \frac{R_4}{R_3 + R_4}.$$

If V was constant, the ratio of

$$\frac{V_x}{V_s}$$

is given by

$$\frac{V_x}{V_s} = \frac{R_4(R_1 + R_2)}{R_2(R_3 + R_4)} = \frac{R_4}{R_2} \text{ for } R_1 + R_2 = R_3 + R_4.$$

Since the method does not involve the flow of current through V_x or V_s the potential difference read is really the *electromotive force,* and not the variable potential difference which a cell gives under any conditions. Since this method enables one to compare the electromotive force of a standard cell against an unknown potential, the potential can be determined in terms of standard electromotive force. With the modern accurate potentiometers such comparisons can be made with a very high degree of precision (i.e., to one part in a million).

(3) For the measurement of very high potential differences and very low potential differences, we make use of a phenomenon which does not fall in the scope of this part of the course. If a potential difference exists between two points, a charged body experiences a force which is the result of the electrostatic field produced. It will be shown in Chapter XVI that this force is proportional to the potential difference under certain conditions. As a matter of fact, by the determination of the absolute potential difference in the electrostatic system, by this means, knowing the relation of the electromagnetic unit of potential and the volt to the electrostatic unit, enables us to have a very accurate determination of potential from the forces

observed. The instrument used for this purpose is what is called the attracted disc, or absolute electrometer. By this device and other static devices based on the principles discussed, potentials ranging from tens of thousands of volts down to thousandths of volts can be accurately measured and compared. It might be added that the use of electrostatic measurements for potentials that are very small together with a knowledge of capacities in the electrostatic system enables us to extend our range of measurements to very feeble electrical currents, that is, to currents of the order of magnitude to 10^{-16} amperes. The discussion of these methods, however, involves a knowledge of static electricity and must be deferred until later.

(4) The measurement of potential by the heating effect developed in a resistance wire in a calorimeter is the absolute electromagnetic method. It has too many sources of error to be very accurate, and is rarely used, in everyday practice, for calorimetry is exceedingly tedious and not very accurate.

44. MEASUREMENT OF RESISTANCE

The measurement of resistance may be achieved as follows: Since it is a derived quantity most of these methods make use of the relation between the current and the voltage across the conductor whose resistance is to be determined. The methods may be listed as follows:

(1) By ammeter and voltmeter.
(2) With a galvanometer and resistance box using the substitution method.
(3) With ammeter and calorimeter by application of Joule's law of heating.
(4) By Wheatstone's bridge, or modifications like the Kelvin bridge for very low resistances.
(5) By potentiometer.
(6) Very high resistance can be measured by means of the rate of discharge of a condenser of high capacity using a ballistic galvanometer as outlined in Chapter XXIII.

(1) In the measurement with the ammeter and the voltmeter one has the ammeter in series with the unknown resistance, and the voltmeter across the ends of the resistance. If the voltmeter has a high enough resistance the current registered by the ammeter flows almost exclusively through the unknown resistance. A division of the voltmeter reading by the ammeter reading should give the resist-

ance. *Caution must be used, however, in this method in order to be sure that the voltmeter has a sufficiently high resistance* so that it does not carry an appreciable part of the current registered by the ammeter.

(2) The second method merely places the unknown resistance and a variable resistance box in series with a galvanometer. When the unknown resistance is in the circuit and the potential is applied, a current will flow through the galvanometer. The adjustment of the galvanometer to read a convenient current by applying a satisfactory potential and suitable galvanometer shunt to the system gives a basis of measurement. If now a known resistance whose magnitude is approximately the same as that of the unknown resistance be placed in the circuit in place of the unknown resistance, the galvanometer will show a different current. The variable resistance must then be varied until the deflection of the galvanometer is the same as with the unknown resistance. The change in reading of the resistance box when balance is obtained added to or subtracted from the known resistance used gives the resistance of the unknown conductor. In general, this method is little used as it is less accurate than the direct comparison methods. The reason for this lies in the range of sensitivity of the galvanometers used.

(3) The application of the calorimeter method, which has been discussed under the derivation of the notion of potential difference in Chapter VI, serves as an adequate means of measuring resistance. For measurement, however, the current must be known quite accurately, for the heating varies as the square of the current while it varies only as the first power of the resistance. The method is rarely resorted to except in particular determinations where other resistance measurements are not easily accessible.

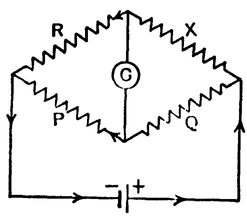


FIG. 45.—Comparison of Resistances by Wheatstone's Bridge.

(4) The most common method of resistance measurements makes use of the Wheatstone bridge. In Chapter VIII, the equation for this method of comparing resistances was derived. In this bridge, the unknown resistance is compared with a known variable resistance by adjusting a pair of known ratio resistances and the variable resistance so that the current through the galvanometer connected in the bridge is 0. In Fig. 45, P and Q are resistances which can be set in definite ratios. X is the unknown resistance and R a standard variable resistance. After having approximately adjusted the resistances P and Q in such a ratio as to give the approximate ratio between R and X, R is varied by small

steps until the galvanometer G gives no deflection. X is then given by

$$X = R\frac{Q}{P}.$$

By making the resistance ratios very accurate and using a sensitive galvanometer resistances can be compared this way to one part in ten thousand without any trouble.

(5) The use of the potentiometer in measuring resistances depends on a knowledge of the current and a measurement of the potential across the resistance. It is superior to the use of the ammeter and voltmeter method in that in this case the potentiometer replaces the voltmeter and one is sure that there is no current flowing through the branch circuit of the potentiometer. Thus, the potential is accurately the potential across the resistance and the current read by the ammeter is accurately the current flowing through the given resistance.

45. POWER MEASUREMENT

Power is measured by two methods: (1) by use of the voltmeter and ammeter; (2) by the wattmeter.

(1) Since the power consumed in watts is the current in amperes multiplied by the potential in volts, the power consumption in a motor can be obtained by multiplying a reading of the voltmeter and a reading of the ammeter. These measurements are applicable strictly only to *direct current* measurements. In the case of alternating currents the voltage and the current may not be in phase, in which case the voltmeter-ammeter readings will give a much higher power reading than the true power consumption in watts. (2) A simultaneous voltmeter-ammeter reading is given by the so-called wattmeter, where the power consumption is given by the current and potential in their proper phase relations. This was adequately discussed in Chapter IX.

46. IMPORTANCE OF CALIBRATION

It may be pointed out that in all these measuring devices except for those methods which have been specifically designated as giving the absolute values of the currents, the instruments used require special calibration in terms of the absolute standards. Thus, every voltmeter, ammeter, wattmeter, resistance box, and galvanometer has either been calibrated in terms of absolute standards or requires calibration before it can be used. For accurate work, these calibra-

tions should be of recent date. If not, they must be checked up before proceeding. The recalibration may be performed at nominal expense either in the testing laboratory by the maker of the instrument in question, or at the National Bureau of Standards in Washington.

47. LEGAL STANDARDS

Before finishing the chapter, it may be well to add a few words concerning the so-called legal standards of current, potential, and resistance by which the measuring instruments of all laboratories may be tested. For such standards, *convenience*, *accuracy*, and *ready reproducibility* are paramount. These standards have been devised by international commissions of engineering and scientific experts, and while being the legal standards depend for their absolute value on measurements made with the methods outlined earlier in the text.

The legal standard of current is based on the rate of deposition of silver from a solution in a carefully defined state and in apparatus of standardized design. The legal ampere is the current which deposits 0.001118 *gram* of silver per second.

The legal standard of potential is based on the use of specially made cells from a class of cells which we will later study. The electromotive force of a given cell should depend entirely on the chemical constituents, the concentration of the substances involved in making up the cells, and the temperature. Consequently, if a cell be made under very carefully specified conditions the E.M.F. which it gives is a good standard of comparison and should be accurate to approximately 1 part in 10,000 at a constant temperature. With such a cell the potentiometer enables us to compare other E.M.F.'s and one therefore has a reproducible standard E.M.F. The legal volt is defined by the potential of the standard Weston cell which is exactly 1.0183 international volts at 20° C.

For resistance, the legal standard consists in the resistance of a column of mercury of specific dimensions. In fact, the legal ohm is the resistance of a *uniform* column of mercury 106.3 cm long having a mass of 14.4521 *grams* at 0° C.

CHAPTER XI

LIQUID CONDUCTORS I, ELECTROLYSIS

48. TYPES OF CONDUCTION AND FARADAY'S LAWS

It was stated in Chapter VII in discussing conductivity that electrical conduction is of two types. The first type of conduction is metallic conduction and depends on the transport of electricity by the very small and mobile electrons. Practically all of the non-metallic conductors depend on the transport of electricity by the atoms or molecules of the substances themselves, the atoms and molecules being charged. This constitutes the second type of conduction. The conductivity caused by carriers of this kind first came to light in a study of solutions. It was later extended to explain the conductivity of most non-metallic solids. In certain crystals conductivity is partly electronic and partly of the molecular type, and the remarkable work of Joffé and of Pohl has given us a completely new insight into the problems of such conductivity.

These men have succeeded in measuring the very feeble conductivity of many non-metallic crystals. In their studies they were able to show the electrolytic nature of the conductivity of some of the crystals, that in certain crystals one ion is mobile while the other ions are fixed and in others that apparently both ions are mobile. Joffé further investigated the influence of temperature on this conductivity and showed that in some crystals temperature acted by decreasing the viscosity of the crystal and also by increasing the dissociation of the crystal lattice into ions thus facilitating conduction. In certain crystals conductivity was found to be influenced by impurities, and in fact the higher conductivities of certain of the crystals were found to be entirely due to the presence of impurities which could be electrolyzed out or separated by fractional crystallization. It was shown that in certain cases, particularly in crystals containing colloidal metals which had been separated by the action of x-rays, the ultraviolet light succeeded in producing a marked conductivity which was due to the liberation of photoelectrons from the colloidal metal particles inside the crystal, the electrons then leading to an electronic

conductivity in a dielectric crystal. The fascinating experiments of Joffé and of Pohl can be found described in Joffé's "Physics of Crystals" and in Gudden's book entitled "Lichtelectrische Erscheinungen."

Most pure liquids, such as benzene, alcohol and water, as well as certain solids such as porcelain and glass when cold, show little conductivity of either type. However, in water when salts, acids or bases are present, the conductivity may become quite high. The same holds true if the glass or porcelain be heated to several hundred degrees.

It was discovered about the same year that Volta developed the voltaic pile that an electric current flowing through water containing traces of dissolved inorganic material caused an evolution of gas at the metallic terminals in the liquid. These terminals, which we will from now on call *electrodes*, were covered with bubbles of gas. It was observed that at the negative electrode hydrogen was liberated, and at the positive electrode oxygen was liberated. It was also observed that twice as much hydrogen by volume was liberated as oxygen. Faraday investigated this and showed that in solutions of other substances such as $CuSO_4$, and $AgNO_3$, the metal was liberated at the negative pole, and that oxygen was liberated at the positive pole, if platinum or non-corrosive electrodes were used. He investigated the deposition of these metals quantitatively and found that for a *given current flowing for a given time the weights of substances liberated were always the same, the weight for each substance being characteristic of that substance.* He further observed that the *weights of these substances liberated had very definite ratios to each other.* He also found *that these ratios were related by simple fractions to the atomic weights of the substances.* He named the phenomenon *electrolysis*, and he called the *negative* electrode *the cathode* and the *positive* electrode the *anode.*

Faraday noted that a quantity 96,500 coulombs (amperes × seconds) of electricity liberated 1 gram atom of hydrogen, $\frac{1}{2}$ gram atom of oxygen, 1 gram atom of silver, $\frac{1}{2}$ gram atom of copper. The same quantity has since then been found to liberate $\frac{1}{2}$ gram atom of zinc, 1 gram atom of chlorine, $\frac{1}{3}$ gram atom of bismuth, and $\frac{1}{4}$ gram atom of thorium. By the term *gram atom* we mean that *weight in grams corresponding to the atomic weight of the substance* as given in any table of atomic weights. Now since 1 gram atom of hydrogen, silver, copper, and so forth, contains, according to Avogadro's rule, the same number of atoms, it follows that for 1 atom of hydrogen there is deposited $\frac{1}{2}$ atom of oxygen, 1 atom of chlorine, 1 atom of silver,

$\frac{1}{2}$ atom of copper, $\frac{1}{2}$ atom of zinc, $\frac{1}{3}$ atom of bismuth and $\frac{1}{4}$ atom of thorium. Thus, if an atom of hydrogen carries a unit charge an atom of oxygen carries two unit charges, that of chlorine one unit charge, that of silver one unit charge, that of copper two unit charges, that of zinc two unit charges, that of bismuth three unit charges, and that of thorium four unit charges.

This phenomenon is shown by the experiment illustrated in Fig. 46. In this figure one sees in series with a battery B and an ammeter A four devices containing different solutions. In the first one, labeled H, we have an H-shaped tube with an electrode in each of the vertical branches of the H. These branches are initially filled to the stopcock at the top with acidulated water. Down the sides of the

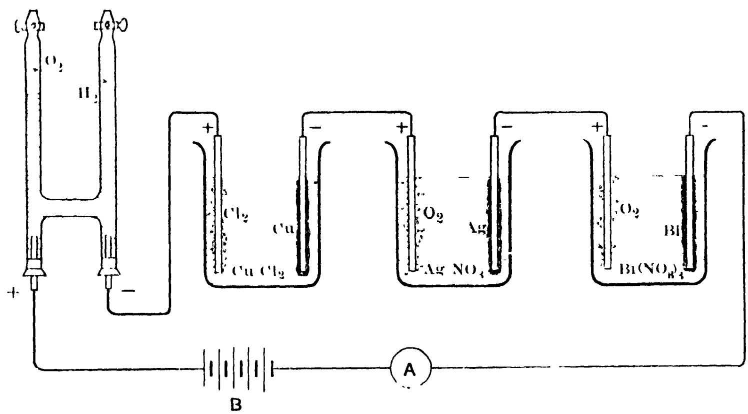


Fig. 46. - Electrolysis.

two vertical tubes there are graduations giving the volume, starting with zero at the stopcock. As the current passes through these cells the positive hydrogen ions go to the negative electrode, give up their charge and are liberated as a gas which rises to the top of the tube. On the left hand side, at the positive electrode, the OH ions which are negatively charged give up their charge to the anode. In doing this they react together to give an atom of oxygen and form water. The atoms of hydrogen and oxygen liberated at the cathode (−) and anode (+), respectively, immediately unite in pairs to form molecules. The molecules then, since the water is saturated with these gases, are liberated in the form of bubbles and the water in both arms of H is displaced by the gas through an outlet tube not shown in the diagram. By adequately adjusting the pressures through this outlet tube the actual volume of gases of hydrogen and oxygen liber-

ated can be estimated at standard temperature and pressure. Thus the weight of gas liberated may be computed and the flow of current determined from Faraday's laws. Experiment would show that the *volume* of oxygen liberated is one-half that of hydrogen. In the beaker labeled $CuCl_2$ there is a solution of cupric chloride, $CuCl_2$. In this the copper is deposited at the negative pole while chlorine gas is liberated at the positive pole. Some of the chlorine will react with water to form hydrochloric acid and liberate oxygen. Some of it will, if possible, react with the anode metal forming a chlorine salt of the anode and finally the balance of the chlorine will be liberated as a gas. If the copper liberated at the cathode be weighed it will be seen that if 1 gram of hydrogen is evolved in the acid tube *H*, there will be 8 grams of oxygen liberated, and $\frac{63}{2}$ or 31.5 grams of copper deposited in the cell Cu. In the beaker $AgNO_3$ there is a solution of silver nitrate. In this beaker silver is deposited on the cathode while the nitrate ion gives up its charge to the anode, and thereupon acts with water to liberate oxygen and form nitric acid. In the silver cell the weight of silver deposited under the conditions that deposited 1 gram hydrogen will be 107.9 grams. In the last cell labeled $Bi(NO_3)_3$, the current deposits bismuth from bismuth nitrate at the cathode, while again the nitrate ion liberates oxygen at the anode. In this case the same current which deposited 1 gram of hydrogen will deposit $\frac{208}{3}$ or 69.3 grams of bismuth.

The metallic *elements* like silver, copper, and so forth, carry *positive charges* for they move to the negative pole, and the *non-metallic* elements like chlorine carry *negative charges* for they move to the positive pole. This peculiar relationship of the charges carried by the various atoms has a definite relation to the combining ratios of the elements in the substances studied. Thus, one hydrogen atom combines with one chlorine atom in hydrochloric acid; two hydrogen atoms combine with one oxygen atom in H_2O; two chlorine atoms combine with one copper atom in $CuCl_2$; three chlorine atoms combine with one bismuth atom, etc. In fact, the charge carried by the ion, that is, the charged atom of matter in electrolysis, is very closely related to the chemical *valency* of the substance, in each compound studied.

Thus, accepting Avogadro's hypothesis and the assumption of Faraday's law that 96,500 coulombs deposit a gram atom equivalent of the particular atom, one is led to conclude that *electricity is atomic in nature.* That is to say, there is a *unit quantity of electricity*, for that is the assumption underlying the reasoning just preceding. At the time of Faraday, this idea was not ripe because people were thinking in terms of the one and two fluid theories of electricity. However, in

1867 Helmholtz stated that the only interpretation to be placed on Faraday's laws was that electricity exists in definite units. In other words, he stated that there was an atom of *electricity* and that it was associated with the *atoms of matter* in *electrolysis*. It was not until thirty years later that J. J. Thomson actually showed the existence of the negative atom of electricity, the so-called electron, and it was nearly ten years later that Millikan made the first accurate measurement of the electron, this unit of electricity in terms of electromagnetic and electrostatic systems of units.

The best way of summing up Faraday's laws would be to state that 96,500 *coulombs of electricity will liberate a gram atom equivalent of an element from solution.* The equivalent can be determined by dividing the atomic weight of the element by the valency of the element in the solution, or the comparison of the amount liberated, with the amount of an atom whose electrochemical equivalent is known which is liberated by the same current. As an example, 96,500 coulombs of electricity liberate 1 gram of hydrogen, 8 grams of oxygen, 107.88 grams of silver, 31.78 grams of copper from $CuCl_2$, 69.3 grams of bismuth from $BiCl_3$, and 35.45 grams of chlorine. This fact is of great use in the measurement and comparison of currents. In fact, these laws furnish the basis of the *legal standard* of current, and the *instrument used in measuring currents by this method* is known as the *voltameter*.

49. CONDUCTION IN SOLUTION AND ATOMIC STRUCTURE

It is now necessary to discuss the nature of this conduction. In the case of metals it was stated that the electron moved freely and probably through nearly forceless channels among the atoms of the crystal lattice. The existence of the electron in the free state is possible only for atoms which have more electrons than they need to give them dynamical stability.

In 1911 the work of Rutherford made it seem probable that atoms were miniature solar systems consisting of a *massive* but minute, central sun or nucleus of positive electricity containing 0.998 or more of the mass of the atoms. The subsequent work of Moseley and Bohr in 1913 led one to conclude that these positive nucleii were surrounded by electrons equal in number to the positive charges, moving in stable orbits of an elliptical or circular nature. The ordinal number of each element in the periodic table fixed the number of positive charges on the nucleus and hence the number of electrons in orbits about it. Thus *H* has 1 positive charge on the nucleus and 1 electron, *He* has 2 positive charges on the nucleus and 2 electrons, *Li* has 3 positive charges and 3 electrons, carbon 4, neon 10, sodium 11, and so on.

According to the new theories of atomic structure originally put forth independently by Gilbert Lewis and Kossel in 1917, the metallic elements are in a more " dynamically " stable state, as far as the orbits of the satellite electrons are concerned, when they have a configuration corresponding to the nearest inert gas in the periodic table. Thus, sodium strives to have the same number of electrons and the same configuration of electrons as the gas neon, which precedes sodium in the table. In attaining this configuration, however, sodium, which has one more positive charge and one more electron, endeavors to lose the electron and become positively charged. Thus metals in general try to attain the stable dynamical configuration of the nearest inert gas. In attaining this state, they will give up their electrons wherever possible either to another atom which will take them up in order to attain its own dynamic stability, or to the lattice of a metal electrode. On the other hand, the so-called non-metallic elements try to gain the dynamic stability of the configuration of the inert gases immediately following them in the periodic table. Thus, it is known that argon has a stable group of 8 electrons in its outer shell; chlorine, the element which immediately precedes it, has a positive charge, one unit less than this, and consequently one less electron. Its stable shell of 8 is incomplete. If possible, it will steal an electron from an atom, like sodium. It then has a negative charge, but it has attained a dynamically stable configuration. Thus we see that there are two types of elements, those that lose electrons and those that gain electrons. Consequently, it is not surprising that in metals we should find free electrons, and that in substances which are non-metallic the electrons should be picked up as rapidly as possible by the atoms. The whole behavior which is exhibited in the formation of the type of polar chemical combinations which ionize in aqueous solution, and thus conduct the current according to Faraday's laws, is due to the striving of atoms to attain the dynamic stability of the nearest inert gas in the Periodic Table at the expense of incurring electrical instability, i.e., the ionic or charged state.

Now, in salts, acids, and bases we have combinations of atoms held together by the charges on the atoms, the atoms which came together to make the groups having lost or absorbed the electrons. That is, the process of combination of sodium and chlorine to form a salt consists in the transfer of the extra electron of sodium to the chlorine atom. The two atoms are therefore dynamically stable and inert chemically, but electrically they are bound together by their charges. If such a molecule as NaCl be placed in solution in a substance like water, which is characterized by having a high dielectric constant (by a high dielec-

tric constant, it will be seen in Chapter XVI, is meant a substance which has the power of reducing the forces between electric charges), the following happens. When a crystal of NaCl is placed in this liquid, the molecules of the liquid immediately act to decrease the electric forces between the atoms of the crystal. In the continuous heat motions which these atoms undergo, the atoms bound together by weakened electrical forces due to the dielectric action of water will be separated. Thus when placed in water, NaCl will very quickly find itself broken up into sodium ions with a positive electrical charge and the chlorine ions with a negative electrical charge. In the case of substances which are divalent in a compound, such as calcium in calcium chloride, this means, in terms of the picture discussed, that the calcium atom is deprived of two electrons, which makes it assume a configuration like argon, and that the two chlorine atoms have each picked up one electron. Accordingly, in solution when calcium chloride is broken up we have a doubly charged positive calcium ion and two negatively charged chlorine ions.

The term applied to this splitting up is called *dissociation.* The theory of electrolytic dissociation was put forth by Arrhenius in 1887 to explain the conductivity of solutions. The fact that NaCl in solution was dissociated in this fashion was confirmed by the fact that in solution dissociated NaCl acts as if it had twice as many molecules or particles present as there are molecules of sodium chloride. In chemical terms this means that the osmotic pressure of NaCl solution is twice as great as the osmotic pressure to be expected from the number of sodium chloride molecules that would otherwise be present. This gives us a fairly clear picture of what is meant by the term ion, and how we explain the theory of electrolytic dissociation.

The existence of ions, however, as was stated earlier in the chapter, is not confined to solutions alone, although they show it most clearly on account of the mobility of the ions present. In fact many, if not most, heterogeneous substances have ions present in them. For instance, glass when it is molten or hot will conduct an electric current, and a very striking experiment can be made in electrolyzing metallic sodium *from sodium nitrate* at 250° through glass into an electric light globe.

In the incandescent light bulb B of Fig. 47 is an incandescent filament F. F emits electrons and is connected to the negative terminal of a 220-volt D.C. supply main. As E the center of the main is at a 0 potential the potential across the filament whose other end is grounded is -110 volts thus lighting the lamp and causing an emission of electrons from the filament. The lamp is well evacuated so that the elec-

trons shoot out from F in all directions. The lower end B of the bulb is immersed in a molten mass of $NaNO_3$ at about 250° C in an iron crucible I. The positive terminal of the 220-volt D.C. line which is 110 volts above the ground E is connected to the crucible I. Na^+ ions are driven by the field from the crucible through the molten $NaNO_3$ to the surface of the glass. The glass, which is also somewhat conducting due to the Na^+ ions which are present because it is a soda glass in which Na^+ ions exist, has the Na^+ ions driven from the outside surface to the inside surface of the bulb. The Na^+ ions arriving on the inside of the glass bulb pick up the electrons from the filament which are driven towards the end of B in the $NaNO_3$ solution by the field. Now $Na^+ + e^- = Na$, where e is an electron and Na is the atom of sodium, so that the ions on arriving on the inside of the bulb are changed to neutral atoms. The neutral atom is volatilized in the vacuum and deposits on the cold upper part of the tube. As Na^+ ions are removed from the glass by the process new Na^+ ions enter the glass from the $NaNO_3$ and the glass remains unchanged. The quantity of Na electrolyzed through the glass in this fashion accurately obeys Faraday's laws of electrolysis, and the method has been proposed as affording an accurate voltameter. The NO_3^- ions go to the iron of the crucible and liberate oxygen or form oxides and nitrates of iron. Had K_2SO_4 been used in place of $NaNO_3$ the K^+ ions would have replaced the Na^+ which came from the glass. The K^+ which enters, the glass,

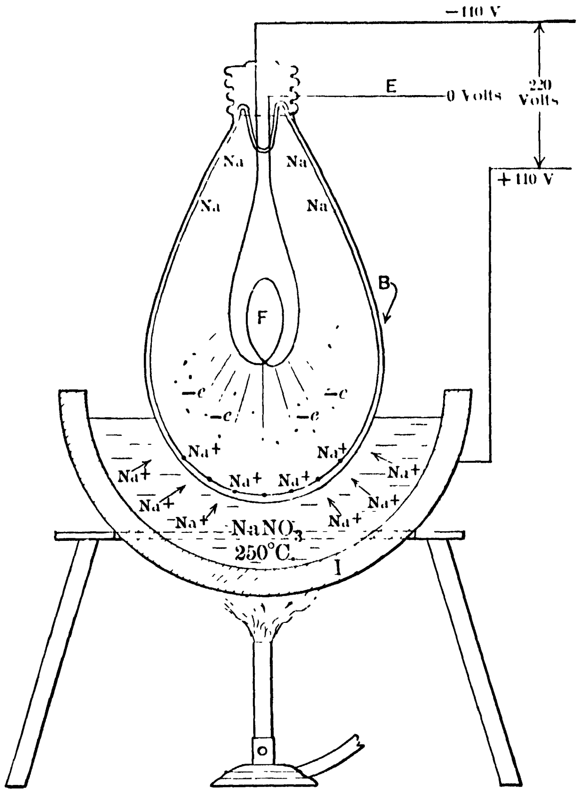


FIG. 47. -Electrolysis of Sodium through Glass.

however, so changes its properties that the glass rapidly goes to pieces.

Another striking effect obtained in certain pure glasses which have a very small number of ions is the complete polarization of the glass in a constant electric field. For instance, Pyrex glass at 400° C will conduct an electric current. If a potential be applied to the glass for 20 minutes, the current will rapidly fall and at the end of 20 minutes the material will act as a good insulator. On reversing the field the current will jump up to its initial value, falling to zero again in the course of time. In this case the conducting ions are due to impurities such as sodium or potassium in the glass, that move to the electrode and owing to mechanical difficulties are not deposited on the surface of the electrode. When these are removed to the negative pole, there are no free ions and the current stops flowing.

The fact that the ions exist in solution has furthermore been verified by a study of the velocities with which the ions move in solution. In fact, we have today quite a definite knowledge of the actual velocity with which an ion of sodium, potassium or hydrogen moves through a solution with a given field strength. The velocity of the ion varies with what is supposed to be the radius of the ion. It is also dependent on the charge of the ion. The larger the ion, the slower its motion. The two fastest ions known are the hydrogen and hydroxyl ions. The high speed of the hydrogen ion is easily understood because a hydrogen ion is a hydrogen atom deprived of its electron. Since the positive *hydrogen nucleus*, called the *proton*, has the smallest volume of any electrical unit known, it is not surprising that it should dodge with considerable freedom among the molecules of the electrolyte. The actual velocities of the hydrogen, sodium, and silver ions in unit electrical field are 0.00323 cm per sec for H, 0.00045 cm per sec for Na, and 0.00058 cm per sec for Ag per unit field.

Electrolysis is used very widely commercially for the following purposes:

(1) For making H_2, Cl_2, O_2, and $HClO$.
(2) For making Na, K, Al.
(3) For the purification of copper and other substances.
(4) For plating of substances with metals.
(5) For the measurement of electrical currents.
(6) It is of great importance through its destructive effects in the corrosion of the frames of steel buildings and other steel structures, for parasitic electrical currents from street cars and power lines daily erode away the metal.

CHAPTER XII

LIQUID CONDUCTORS II, BATTERIES

We now come to a new problem in the question of conductivity f solutions. The subject which we are going to discuss will require he introduction of two new concepts, the one *solution tension*, the ther *solution pressure*.

50. SOLUTION TENSION

If we take a piece of metal like zinc or magnesium and place it in vater, it has a tendency to go into solution. As was intimated in he last chapter, when an electropositive metal like zinc goes into olution it tends to do so in the form of an ion; that is, it tends to go nto solution as an atom which has lost electrons in order to reach ts most stable chemical or dynamical form. If then we place a piece f zinc in water, zinc ions with a double positive charge go into solu- ion. As they leave the metal with a positive charge the metal which vas neutral must retain an equivalent negative charge (i.e., the alence electrons of the zinc which were sacrificed to gain dynamic tability). As solution continues the negative charge on the zinc ccumulates. The negative charge however acts to draw back the ositive ions of zinc and cause them to return to the metal. The endency to go into solution as ions is called solution tension.

51. SOLUTION PRESSURE

The second concept which we must use is that of solution pres- ure. If we have a substance like zinc chloride in solution the zinc ons which are present will tend to condense out on any surface. This phenomenon we call solution pressure. As the ions start to con- lense they carry a positive charge to the electrode and so very quickly harge it up to a potential such that no more positive zinc ions can each the electrode.

52. EQUILIBRIUM

Consequently there are two opposing tendencies when any metal s immersed in a solution of its salts. On the one hand it sends ions

into solution, and on the other hand ions try to precipitate out. These two processes lead to *equilibrium* with a certain number of ions in solution carrying their charge, while the metal has an equivalent number of opposite charges. This equilibrium is determined by the nature of the metal and the form in which it ionizes, the charge on the metal which pulls the ions back, and the concentration of the ions in solution which also limits solution.

If then we place a piece of zinc into a solution of zinc chloride or zinc sulfate, the zinc will go into solution until the concentration of zinc ions produces equilibrium, that is, there will be a solution of the zinc until the zinc acquires a negative charge so that the number of zinc ions striking the electrode per second equals the number that leave the electrode. In this state the zinc electrode has a negative charge. Had we used magnesium instead of zinc the tendency to go into solution would be even greater, and the negative charge necessary to draw in enough magnesium ions to make the rate of precipitation equal the rate of solution would be higher than with the zinc.

In practice the equilibrium potential can be determined for any metal and the metals may then be arranged in the order of the potential which they give under standard conditions. Thus each metal would be characterized by a specific potential with regard to the solution it was placed in, and one would have a series of the solution tensions of the various metals. If one listed the various metals in such a scheme, one would find that magnesium has a very high solution tension, zinc a slightly lower one, hydrogen still lower, and copper and platinum will have still lower tensions. Such a table of elements is seen below.

Voltaic or Replacement Series of the Ions of the Elements Derived from the Heats of Formation as Given by Ostwald

Positive Ions					Negative Ions
Li^{+}	Ca^{++}	Zn^{++}	Sn^{++}	Cu^{+}	OH^{-}
Rb^{+}	Mg^{++}	Fe^{+++}	Pb^{++}	Hg^{+}	Cl^{-}
K^{+}	Al^{+++}	Co^{++}	H^{+}	Ag^{+}	Br^{-}
Na^{+}	NH_4^{+}	Cd^{++}	Fe^{++}		I^{-}
Sr^{++}	Mn^{++}	Ni^{++}	Cu^{++}		$S^{=}$

It must be noted, however, that the table given above is based on the relative heats of formation of the particular ions listed and can only be used after a careful consideration of the conditions pres-

ent in the solutions in which the metals are used to be sure that the reaction is such as is indicated by the table. If certain of the substances can as a result of the chemical reactions possible in a given solution react in a different manner from the one indicated in the table the whole order of the reaction may be modified. For example, Fe in slightly acid solution tends to go into solution replacing hydrogen which as will be seen later is given off at the surface of a copper electrode. If the solution is neutral or alkaline the tendency is much decreased on account of the formation of insoluble hydrated iron oxides. The iron electrode will therefore be negative while the copper electrode will be positive due to the deposition of H^+ ions on its surface. If the acid solution be replaced by an alkaline solution of KCN the reaction is completely changed. Copper will dissolve, giving a complex copper cyanide ion; $Cu(CN)_2{}^-$ as copper in this solution is very soluble. On the other hand in an alkaline solution the iron will show little tendency to replace hydrogen ions. Thus the copper will go into solution as Cu^+ uniting with $2CN^-$ to give $Cu(CN)_2{}^-$ and will leave a negative charge on the copper electrode while the K^+ ions will give a positive charge to the iron electrode and subsequently react to form KOH liberating H_2 gas at the iron electrode. The relative potentials of the Cu and the iron electrodes will thus be reversed from those in the acid solution as will also be the case for the course of the reaction. The reversal can be readily seen by the direction of the deflection of a galvanometer in a circuit in which the Cu and Fe electrodes are first placed in acid and then washed and placed in the cyanide solution.

Substances such as chlorine or oxygen also go into solution. They, however, go into solution taking on a negative charge, that is, they yield negative ions. This results from the fact that both the chlorine atom and the oxygen atom are more stable chemically and dynamically when they have gained enough electrons to fill out their shell of eight, even when in doing so they gain a negative charge. Experimentally, it is difficult to obtain electrodes of gases because gases are non-conductors. The way in which a chlorine gas electrode can be achieved is by bubbling the gas through meshes of a platinum gauze which constitutes the electrode. The adsorbed chlorine then serves as the electrode surface. As it ionizes with a negative charge the electrode acquires a positive charge and this process goes on until the opposing tendencies of solution tension and solution pressure again establish equilibrium. Accordingly in the case of the negative ions again we have a series of elements which exhibit varying tendencies to go into solution. Those which give the electrodes the

highest positive charge are the so-called most electronegative gases, that is, they have the highest electron affinity. Thus chlorine is far more electronegative than oxygen and oxygen very much more so than nitrogen, which really shows no tendency to go into solution in the form of ions.

These principles which we have discussed underlie the whole action of the chemical cells or batteries. We will discuss several different types of cells in what follows.

53. THE CONCENTRATION CELL

Perhaps the simplest of all cells is the *concentration cell.* Suppose we place a zinc electrode in a solution of concentrated zinc sulfate, place another electrode in a solution of very dilute zinc sulfate and then connect the two solutions by means of a conducting bridge filled with zinc sulfate. A cell of this sort could perhaps most easily be achieved by placing the dilute zinc sulfate in a porous porcelain cup and placing this cup in a concentrated solution of zinc sulfate which contains the other electrode. The zinc in the concentrated solution will tend to dissolve, charging this electrode negatively. It will do the same in the dilute solution, but owing to the lower concentration in the dilute solution the equilibrium charge of the zinc electrode in the dilute solution will be more highly negative than in the case of the concentrated solution. One thus has two electrodes of different negative potentials, the potential of the zinc in the concentrated solution being less negative, or consequently more positive, than the electrode in the dilute solution. There is thus a potential difference between the two electrodes.

If a galvanometer be connected to the two zinc terminals current will be observed to flow from the zinc in the concentrated solution to the zinc in the dilute solution trying to equalize the potential of the two electrodes. If one accept the electron as the carrier of the current, the superfluous electrons on the electrode in the dilute zinc solution would really flow to the zinc in the more concentrated solution. In any case, the equalization of charges would take place. As soon as the charge of the zinc in the dilute solution is lowered, more zinc will go into solution, and the flow of negative electricity to the zinc in the concentrated solution will increase the negative charge of that electrode. This increase immediately acts to draw in more zinc ions because the equilibrium potential has been destroyed. Thus, on the concentrated side zinc ions will go to the electrode and give up their positive charge, becoming zinc atoms, while on the negative

side zinc will go into solution. In the meanwhile, current flows through the galvanometer. It is seen at once that the direction of the processes at work are such as to reduce the concentration of the strong solution and increase the concentration of the weak solution, the net result being the equalization of the concentrations of the two solutions. Thus we keep the electrical current flowing as long as the concentrations are different. The energy which is used during this process corresponds to the heat of dilution of the zinc, for what has occurred is that the zinc which was in a concentrated solution is now occupying the total solution volume of both the concentrated and the dilute solution. In studying the energetics of such cells, the electromotive forces would at first sight appear to be computable from the heats of dilution alone. This is so in some cells, but not in others. In general another term must be regarded, and that is the *entropy* of the reaction. A study of this lies beyond the scope of this course and belongs properly in an advanced course in physical chemistry.

54. THE COPPER CELL OR DANIELL CELL

One may now discuss the case of a cell made of dissimilar metals. Place a dilute solution of zinc sulfate inside of a porous cup, place the porous cup in a concentrated solution of copper sulfate. Into this copper sulfate solution may be placed an electrode of copper. Now, as before, the zinc takes on a negative charge, zinc ions going into solution. In the case of the copper equilibrium is set up by a few copper ions going into solution or perhaps if the solution is concentrated enough a few copper ions precipitating out. In any case the potential of the copper relative to the potential of the zinc is distinctly more positive. The instant a wire is attached to the two electrodes making a transfer of electricity from one to the other possible, the transfer occurs. The equilibrium which had existed is at once disturbed. The copper which was before in equilibrium, owing to the loss of the positive electricity which flowed to the very negative zinc electrode, now strongly attracts the copper ions. The copper is deposited rapidly and one finds the electrode coated with small nodules of new metallic copper. In the case of the zinc electrode the zinc goes into solution as soon as the equilibrium is disturbed because through the advent of the positive electricity the potential of the zinc electrode is not sufficient to hold back the zinc ions. Thus the action of this cell consists in the solution of zinc and the precipitation of copper. The laws of electrolysis lead us to suppose that as many ions of zinc go into solution as ions of copper go out of solution

for they are both bi-valent. In this case, however, the chemical action would go on until all the zinc was dissolved or copper was liberated. The energy which is expended by the current then comes from the replacement in solution of copper ions by zinc ions, that is, it corresponds to the differences in heat of solution between zinc and copper, with of course a correction for the entropy term involved as mentioned under the concentration cell.

55. THE ACID CELL

It has been mentioned in the first sections that the various elements could be placed in a series as regards the potentials which they attained. This corresponds to the electromotive force series of the elements and was first discovered by Volta in his study of the voltaic pile. One way of investigating the replaceability would be to place the metal to be tested into a solution of the metal which it is expected to replace. That is, if one place a piece of iron which has a higher solution tension than copper into a solution of copper it will be observed that the iron goes into solution and the copper goes out of solution. The action is very much like that of the Daniell cell, the precipitation of copper taking place at points where there are bits of impurity which decrease the solution tension of the iron. Again if a piece of zinc be placed in an acid solution hydrogen is liberated as a gas, that is, hydrogen has a lower solution tension than zinc and is thus replaced by it. The relative solution tension may be measured in three different ways: (1) by the method of replacement just discussed, (2) by the relative potentials acquired by the metals in solution, and (3) by the relative energy liberated by the elements when they go into solution.

In discussing this table of solution tensions we find that hydrogen stands higher in the series than either platinum or copper. Thus, hydrogen replaces copper in solution, and tends to go into solution taking on a positive charge more readily than copper does. It is well known that copper will not dissolve in HCl. This fact enables us to explain the action of the third type of cell, the *acid cell.*

Suppose a zinc electrode and an electrode of copper be placed in a solution of hydrochloric acid. As before, the zinc atoms go into solution charging the zinc negatively. It is possible that a small number of the copper ions may go into solution having a very low solution tension, but this number will not be great. In the solution besides the ions of zinc and the few ions of copper there are *many hydrogen ions.* These ions will attempt to precipitate out because hydrogen

has a solution pressure. On immersion in the acid the copper electrode will come to equilibrium with the hydrogen ions as well as with the copper ions. If, now, the copper electrode and the zinc electrode be connected, the negative electricity on the zinc electrode will flow to the copper. The potential will then become too negative for the equilibrium of that electrode with the solution. At first, the copper having the lowest solution tension will precipitate out. The potential of the electrode will however be even more negative than will permit the existence of hydrogen ions in the solution in such a concentration, and the hydrogen ions will proceed to precipitate out. Meanwhile, the zinc ions continue to go into solution giving more negative charges to the copper electrode, and thus more hydrogen ions are electrolyzed out. As soon as hydrogen ions reach the electrode, they give up their charge and go off as bubbles of hydrogen gas. Their positive charge draws more electrons from the zinc electrode and consequently more zinc can go into solution. Thus the zinc goes into solution and the hydrogen goes out of solution. Since the hydrogen has one positive charge, two hydrogen ions will go out of solution for every zinc ion that goes into solution. The current will flow until all the zinc has gone into solution or all the hydrogen ions of the acid have been replaced by zinc ions. The energy in this case is the difference in the solution energy of zinc ions and hydrogen ions, that is, the energy given by this cell is the energy of formation of zinc chloride $ZnCl_2$ less the energy of formation of hydrochloric acid HCl, with the added restriction caused by the neglect of the entropy term.

56. POLARIZATION

The acid cell is the one most frequently encountered in chemical reactions. Practically, however, it is not useful because of the phenomenon known as polarization. This phenomenon of polarization is due to the liberation of hydrogen gas. It is obvious that as soon as the hydrogen is liberated it forms bubbles. These can make a more or less continuous coating on the electrode which does not conduct electricity. Such a coating reduces the area available for the passage of the current. It increases the internal resistance of the cell, and thus enables the discharge of the electrodes to take place more rapidly than the charge. The potential difference that will be maintained between the two terminals due to the rapid discharging of the terminals and their slow rate of charging because of the high resistance caused by reduced surface area, will therefore not be the potential difference at which one could expect the cell to operate in the absence

of polarization. As a number of convenient cells used in commercial practice have a tendency to suffer from polarization, features have been arranged to obviate this. All that is required for the removal of polarization is a removal of the gaseous film by: (1) the mechanical jarring or scraping of the electrode to remove gas bubbles, (2) a chemical method depending on the use of oxidizing agents such as K_2CrO_4 or MnO_2, (3) the use of solutions which do not polarize. The Daniell cell discussed in case No. 2 is a non-polarizing cell inasmuch as copper is deposited and this forms a conducting film.

57. CALCULATION OF THE APPROXIMATE E.M.F. OF A CELL

We now take up another important phase of the question of cells, that is, the value of the E.M.F. given by a cell. As was stated in Chapter VII, the internal resistance of the cell prevents current being supplied to the electrodes as fast as it is removed by the connecting leads. Thus the potential difference given by a cell with a resistance across it is not the maximum potential difference which the cell can give. The maximum potential difference which any cell can give is called the E.M.F. of the cell and can be measured by measuring the separate potentials of the two electrodes relative to the solution or by measuring the actual potential difference between the two electrodes; using a method which does not consume current. Such a method would include the use of a potentiometer or the use of an electrostatic voltmeter.

The E.M.F. of a cell may be approximately computed by a method that was hinted at in a discussion of the source of energy of the cell. Potential difference, or E.M.F., is by definition equal to the work involved in the transport of electricity from one terminal of a cell to another divided by the quantity of electricity transported. That is, $W = Eq$, where W is in ergs if q is in absolute units of quantity and E is in absolute units of potential.

In the last chapter we found that 9650 absolute units of quantity liberate (atomic weight in grams)/valency of the substance. Let us designate the number of grams liberated by 1 absolute unit as Y. Accordingly $Y = \dfrac{\text{atomic weight}}{9650 \times \text{valency}}$. Then q absolute units will liberate qY grams. If the qY grams of the substance instead of being liberated had been consumed in the reaction giving the current in the cell, they would also have furnished q absolute units of electricity, provided we neglect the consumption of the substances which go to make up the energy involved in the change in entropy of the reaction.

Let us neglect the entropy. If each gram of substance consumed furnished H calories, where H is the heat of chemical reaction per gram of substance transformed, then this would have used qYH calories or $qYHJ$ ergs of work. This work must be equal to the potential difference E times the quantity q transferred, for the electrical energy transported comes from the chemical energy consumed. Thus we can write $qYHJ = Eq$, or $YHJ = E$. The E here obtained is in absolute E.M.U. of potential, and to obtain E in volts we must divide by 10^8. Hence we can write $E_{\text{volts}} = \frac{HYJ}{10^8}$.

Now we do not usually employ the quantity Y as the expression for the deposition of the substance. In place of it we use the *electrochemical equivalent* S, where $S = \frac{\text{atomic weight in grams}}{96{,}500 \times \text{valency}} = \frac{Y}{10}$, i.e., S is the mass of substance liberated per coulomb. If S is used $E_{\text{volts}} = \frac{HSJ}{10^7}$. We are thus in a position to calculate the E.M.F. of any cell in volts if the quantity H is known and if we can neglect the entropy change. H is usually the difference between the heat of formation per gram of the substance going into solution and the substance going out of solution.

Example.—In the case of the oxygen-hydrogen cell we can calculate the approximate E.M.F. with considerable ease. The heat given out by the reaction is equal to 34,000 calories per gram of H consumed. J is equal to about 4.2×10^7. S for hydrogen is 0.00001044. Dividing the product of these quantities by 10^7 to convert them into volts E comes out as 1.49 volts. In a similar fashion we can calculate the E.M.F. of any cell, if we know the necessary quantities, and can omit the entropy factor. It is to be noted that in such calculations the value of S for the reaction must apply to the same element as the one for which H is calculated.

58. CAUTIONS IN ELECTROPLATING

There is one point of great importance which the question of solution tension brings out in regard to the electroplating of metals. As was stated in the last chapter, the passage of the electric current causes the deposition of the ion to take place. It must be borne in mind, however, that in order to cause zinc to deposit from a zinc sulfate solution the potential of the electrode upon which the zinc is being deposited must be more negative than the potential which a zinc electrode immersed in the solution would take on due to the tendency of the zinc to go

into solution. That is, in any electroplating bath we must be careful to have the potential sufficiently great to insure the substance being deposited. Thus, if we have a solution containing a bismuth electrode which is to be plated with copper, the copper having a higher solution tension than bismuth, we must remember that Bi and Cu form a cell which has an E.M.F. tending to cause Cu to dissolve and to deposit Bi. Thus to electroplate the Bi with Cu we must have a P.D. greater than the E.M.F. of the Bi-Cu cell. It is also useful to know that if one, for instance, wishes to plate an electrode with a given metal and one has present metals which are more readily deposited than that metal the other metal will also tend to precipitate out when the potential is applied to the bath. One mode of separation of copper in a pure form is by means of electrolysis with carefully controlled potentials so that the copper deposits whereas other more electropositive impurities do not. Furthermore the use of too high a potential and hence too great a current density causes rapid irregular deposition and does not give an even uniform deposit.

59. STORAGE CELLS

The peculiar reversibility of a cell, namely, the fact that copper and zinc in the proper solutions will, if connected, deposit copper and dissolve zinc, giving us an electrical current, and if a sufficiently high potential be applied to the cell to reverse the action, the cell will consume electricity and deposit zinc putting copper into solution, leads us at once to the principle of the storage battery. The storage battery is merely a convenient reversible form of cell. In such a cell when we permit the electricity to flow into the cell we get a deposition of the element having the greatest solution tension. When all this element has been deposited on the proper electrode we may then by connecting this electrode to the other electrode reverse the reaction and obtain current at the expense of the chemical energy stored up. In order to be able to have the greatest capacity and potential with as few cells as possible, and in as convenient a form as possible for practical purposes, a great many investigations have been made. There seem to be two types of cells on the market which are about equally successful. The first of these is the lead storage battery and the second is the Edison cell.

In the lead storage battery, the electrodes consist of grids in a framework of lead filled with lead as one electrode and lead peroxide made into a paste with lead sulfate as the other electrode. The elements of the battery consist then of electrodes of lead and lead

peroxide, PbO_2. The mechanism of the reaction is summed up in the following equation:

$$PbO_2 + 2H_2SO_4 \rightarrow Pb(SO_4)_2 + 2H_2O \rightarrow Pb^{++}SO_4 + 2H_2O + SO_4^{=}.$$

That is, PbO_2 reacts to form $Pb(SO_4)_2$ which ionizes, sending $SO_4^{=}$ ions into solution, leaving the PbO_2 plate positively charged and filled with $PbSO_4$. As $PbSO_4$ is insoluble there is little Pb in ionized form in the solution and the ions depositing on the Pb plate are other ions. As the solution contains an excess of $SO_4^{=}$ ions these tend to deposit on the lead plate charging it negatively. On closing the circuit between the PbO_2 plate and the Pb plate equilibrium is disturbed causing more $Pb(SO_4)_2$ to ionize and causing more $SO_4^{=}$ ions to deposit on the Pb plate. These ions on losing their charge will react with water to give O_2, which will in part react to oxidize the Pb plate to PbO, and in part will come off as a gas, depending on the rate of discharge of the battery. The action will continue until all the PbO_2 is changed to PbO. Under these conditions the cell is said to be discharged. If, now, an external electromotive force be applied to the cell, the positive terminal being on the PbO_2 electrode and the negative terminal on the PbO electrode, and the electromotive force exceeds that of the storage cell, the following reaction will take place: the acidulated water will furnish hydrogen ions which will at once migrate to what was initially the lead plate if the negative potential is great enough to cause hydrogen ions to deposit. If the hydrogen ions give up their charge, atomic hydrogen is liberated. This reacts with the oxygen on the lead electrode, giving water and leaving pure lead. At the positive and what was initially the PbO_2 electrode $SO_4^{=}$ ions will be attracted. Arriving at the electrode they will give up their charge and react with the water liberating atomic oxygen. This oxygen will at once react with the PbO in the form of $PbSO_4$ giving PbO_2. After all the PbO of the positive plate has been converted to PbO_2 the cell will be in its initial state, and is said to be charged.

The cells can act in this fashion indefinitely until mechanical wear and tear destroy them. To reduce the internal resistance and increase the capacity of the cell it is necessary to have the plates as large as possible in area and as close together as possible. Contact between the plates is prevented by thin porous separators. If one portion of a plate is closer to its opposite plate than the other the current flow will be concentrated at this point. The result will be that the plate will wear or erode unequally. Thus, buckling of the plates

leads to very rapid destruction of the cell. If the cell be very rapidly charged or discharged, the mechanical action of the gases involved as well as the heat produced by the resistance of the cell leads to buckling of the cell plates and to rapid destruction. The evolution of gases also serves to mechanically loosen the paste in the grids. Owing to the fact that in the process of charging the water is electrolized and owing to the fact that gradual evaporation takes place, it is necessary that the batteries have their solution replenished by the addition of fresh *distilled* water from time to time. If a battery remains unused for long periods of time there is a gradual solution of the lead and precipitation of lead sulfate at the bottom of the cell. This process destroys the cell and removes sulfuric acid. All cells have instructions giving the percentage of acid present when charged and when discharged in terms of the specific gravity of the solution and the cell should be tested both when charged and discharged by measuring its specific gravity. The efficiency of the lead storage cell, if properly handled, is from 75 to 85 per cent.

The Edison cell, which consists of a hydrated nickel oxide electrode on the one hand and an iron electrode on the other, is far more mechanically robust. On discharging, the iron oxidizes and the nickel oxide is reduced. The cell is one using alkali as an electrolyte instead of acid. Owing to the lower specific gravity of iron and nickel it weighs half as much as the lead storage cell for the same capacity. On the other hand, its E.M.F. is only 1.25 volts, while the E.M.F. of the lead cell is about 2.2 volts. This means that while the Edison cell weighs half as much as the lead storage cell nearly twice as many cells are required to give the same potential. The efficiency of the Edison cell is only 50 per cent. Deterioration of the Edison cell, however, is far less and it can stand much mishandling.

60. DRY CELLS

Besides storage cells, the only other cells which are of particular interest today are the zinc carbon cells or modifications of them which we know as "dry batteries." These batteries consist of a zinc electrode surrounding the cell with a central core of graphite surrounded by a paste of moist MnO_2. These cells are polarizing cells and their action is very much like the action of the acid cell. Zinc goes into solution at the negative pole and hydrogen is liberated on the carbon; in this case the aqueous solution is replaced by a moist paste of gypsum with ammonium chloride solution. The manganese dioxide reduces the polarization, but heavy currents cannot be

drawn from the batteries for any length of time without a fall in potential.

61. STANDARD CELL

Standard cells are cells which are used as standards of comparison of electromotive forces. Many types are on the market and they are so designed as to give an accurately reproducible and constant electromotive force when made up according to formula. The Weston standard cell consists of an H-shaped tube, as shown in Fig. 48. At one lower end of the *H* a platinum wire is sealed into the tube. This makes contact with the solution by being covered with an amalgam of cadmium and mercury. Next to this amalgam is a saturated solution of cadmium sulfate, a few crystals of the latter lying on the amalgam. On the other side of the *H*, another platinum wire is sealed in and this is covered by a mercury globule. Over the mercury is a paste of mercurous sulfate. The cross part of the *H* is filled with a saturated solution of cadmium sulfate. The two upper arms of the *H* are sealed off. The E.M.F. given by this cell should be 1.0183 volts at 20° C. The E.M.F. of all cells varies with the temperature, and standard cells should always be used at their rated temperature.

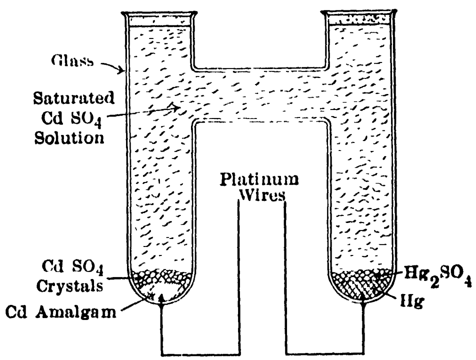


FIG. 48.—Weston Standard Cell.

62. MOST EFFICIENT ARRANGEMENT OF CELLS

In discussing cells one might at this point discuss the question of the most efficient use of cells when they are used in quantity. There are two different ways of using cells in a circuit, adapted to two extreme types of usage. In the first place, cells may be connected in parallel. The use of cells connected in parallel is of importance where a heavy current is required and the external resistance R_e is less than the internal resistance R_i. Assume the cells have all the same E.M.F. equal to E. Since the cells are all in parallel, they pass the current i through the resistance. The equation then is

$$E = i(R_{\text{cells}} + R_e).$$

In this equation, R_{cells} is the total resistance of the n cells in parallel. Since

$$\frac{1}{R_{cells}} = \frac{1}{R_\iota} + \frac{1}{R_\iota} + \frac{1}{R_\iota} \cdots = \frac{n}{R_\iota},$$

therefore,

$$E = i\left(\frac{R_\iota}{n} + R_e\right).$$

Thus the current

$$i = \frac{E}{\dfrac{R_\iota}{n} + R_e}.$$

If R_e is small compared to R_ι, the current i will then be proportional to n the number of cells in parallel.

Again suppose the cells to be placed in series. The arrangement

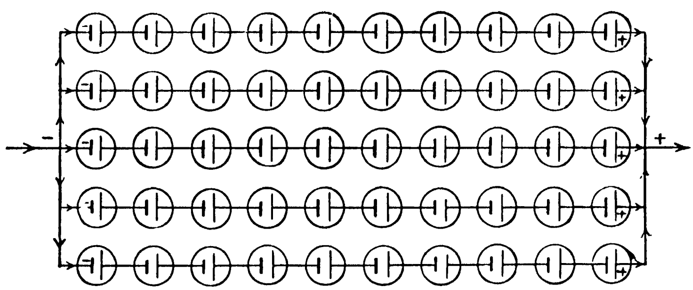


FIG. 49.—A Bank of 50 Cells Arranged with 10 Cells in Series and 5 Rows of Cells in Parallel.

of cells in this form is of special use for the case where the external resistance is great compared to the internal resistance R_ι of the cells. If there are n cells

$$E_{total} = nE = i(nR_\iota + R_e),$$

hence,

$$i = \frac{nE}{nR_\iota + R_e}.$$

If R_e is very great compared to R_ι, i is proportional to n.

In general, where many cells are to be used and a specific circuit is to be run in the most efficient manner, that is, to get the greatest current, one cannot offhand tell what the best arrangement of the cells will be. Assume that there is a total of n cells. Suppose in the

most efficient arrangement which we intend to calculate, that x groups of cells should be in parallel. $\frac{n}{x}$ cells will then be in series for $x\left(\frac{n}{x}\right) = n$, by the geometry of series and parallel connections. This relation is illustrated in Fig. 49 where $x = 5$, $n = 50$, whence $\frac{n}{x} = 10$ cells are in series. The current in this case will be expressed by an equation

$$i = \frac{\frac{n}{x}E}{\frac{n}{x}\frac{R_i}{x} + R_e} = \frac{E}{\frac{R_i}{x} + \frac{x}{n}R_e}.$$

This equation may at once be set up from what has gone before concerning the current when cells are in series and in parallel.

Now i is to be a maximum. Calculus has taught us that we can determine for what value of one of the quantities an equation is a maximum by differentiating the quantity with respect to the variable which is to determine whether it is a maximum or not. In setting the derivative equal to 0 the solution of the resulting equation at once gives us the condition which must be fulfilled to give the maximum current. Thus the maximum efficiency, or perhaps a minimum efficiency, is found by setting.

$$\frac{di}{dx} = 0$$

$$\frac{di}{dx} = 0 = \frac{-E\left(\frac{-R_i}{x^2} + \frac{1}{n}R_e\right)}{\left(\frac{R_i}{x} + \frac{x}{n}R_e\right)^2}.$$

Solution gives us the result that

$$\frac{R_i}{x^2} = \frac{1}{n}R_e,$$

or that

$$\frac{R_e}{R_i} = \frac{n}{x^2} = \frac{1}{x}\left(\frac{n}{x}\right).$$

To test whether this is a maximum or a minimum the second derivative of the expression must be taken and the sign of the result must be observed. It is hardly necessary to carry this out here, and it suf-

ices to state that the condition laid down is the condition for a naximum current. The relation

$$\frac{R_e}{R_i} = \frac{1}{x}\left(\frac{n}{x}\right)$$

n words reads as follows:

$$\frac{\text{number of cells in series } \left(\frac{n}{x}\right)}{\text{number of cells in parallel } (x)} = \text{the ratio of the external resistance to the internal resistance.}$$

f the two resistances are equal (i.e., if $R_e = R_i$, the number of cells 1 series equal the number in parallel).

Example. To see the general application of the problem assume ne has 100 dry cells to connect so as to give the maximum current. Assume that the external resistance is 10 ohms and that the internal esistance of a cell is 2 ohms. The rule says that:

$$\frac{\text{number of cells in series}}{\text{number of cells in parallel}} = \frac{10}{2} = 5 = \frac{1}{x}\left(\frac{n}{x}\right).$$

That is, consequently, $5x = \frac{n}{x}$. Therefore $x^2 = \frac{100}{5}$. Since $x = \sqrt{20} = 4.48$, the result says that there should be 4.48 cells in parallel nd consequently $\frac{100}{4.48}$ cells in series. Actually one cannot split up ells into fractions. *We consequently have to take the closest whole umber to* 4.48 *cells and this will be the most efficient way in which the ells can be combined.* Thus, if we place 4 cells in parallel and 25 in eries we should get the maximum current. Since 4.48 is so nearly qual to 4.5 this indicates that we should get almost as much current f we placed 5 cells in parallel and 20 in series. To prove this we alculate the current in the two cases. Substituting the values for and $\frac{n}{x}$ as well as the resistances in the equations, one has for the cells in parallel

$$i = \frac{25E}{25(\frac{2}{4}) + 10} = 1.11E.$$

n the other case one has

$$i = \frac{20E}{20(\frac{2}{5}) + 10} = 1.10E.$$

The difference is exceedingly small. If, however, we took the case of 2 cells in parallel we would see that the current was distinctly less. In this case

$$i = \frac{50E}{50(\frac{2}{2}) + 10} = 0.835E.$$

Had we chosen 10 cells in a series and 10 in parallel

$$i = \frac{10E}{10(\frac{2}{10}) + 10} = 0.835E.$$

Such calculations are often needed in electrical circuits to get the best combinations of cells.

CHAPTER XIII

THERMOELECTRICITY

63. THE PELTIER AND SEEBECK EFFECTS

IN 1826 Seebeck observed a curious effect. He found that if one took a strip of antimony wire and connected its ends to two strips of bismuth, whose ends were connected to the terminals of a sensitive galvanometer, when one of the bismuth-antimony junctions was heated to a higher temperature than the other, the galvanometer showed a deflection indicating a current. This became known as the Seebeck effect, and is illustrated in Fig. 50.

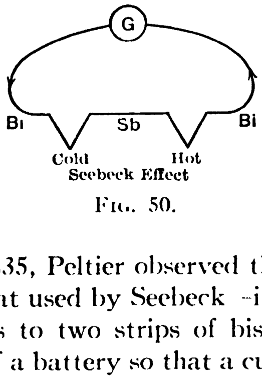


FIG. 50.

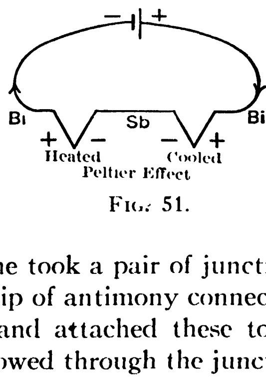


FIG. 51.

In 1835, Peltier observed that if one took a pair of junctions similar to that used by Seebeck—i.e., a strip of antimony connected at its two ends to two strips of bismuth—and attached these to the terminals of a battery so that a current flowed through the junctions, one observed in the case of the junction where the current went from bismuth to antimony a cooling of the junction, and where the current went from antimony to bismuth a heating of the junction. This was called the Peltier effect and is illustrated in Fig. 51.

The two effects were soon discovered to have a common origin and today we give the following simple explanation of the phenomena.

The explanation followed what was originally supposed to be the correct interpretation of metallic conduction. It was assumed that the metals bismuth and antimony had free electrons present, but that bismuth had more free electrons than antimony. At a junction between the two, owing to the higher concentration of free electrons

in bismuth, a diffusion of these from bismuth to antimony took place. Thus, the bismuth metal was left positive and the antimony became negative. The electrification was, however, confined to the actual junction of the two metals.

Now it was further believed that if one heated such a junction, owing to the increased temperature agitation, the bismuth free electrons were able in greater number to traverse the boundary into the antimony. Thus, with two junctions, the hotter one would have the bismuth-antimony potential greater than the cold junction. The net result would be that the current would flow from the hot bismuth-antimony to the cool antimony-bismuth through the circuit.

The explanation of the Peltier effect is also based on the existence of these potential differences at the boundaries of the two metals. When a current is passed across the junction from bismuth to antimony, the bismuth being positive and antimony negative, by the conventional laws of flow of current, positive electricity flows from the bismuth to the antimony. This tends to annihilate the electrical field at the junction. To build up the field again takes heat energy from the metal, consequently heat is consumed and the bismuth-antimony junction is cooled. In the case of the antimony-bismuth pair where the current flows from antimony to bismuth the current must flow from negative to positive in the field produced between the two metals. Electricity is therefore carried across the junction against the field. Thus work is done and the antimony-bismuth couple is heated. The two effects should be about equal in magnitude provided the temperature differences are not too great. This explanation, although initially accepted, has had considerable doubt cast on it in view of our present-day uncertainty as to the exact mode of transfer of electricity in metals. That free electrons are present in small numbers is possible. It is now doubtful whether the free electrons in large numbers are present at all, at least in a state where the electrons can share in the temperature equilibrium of the atoms. In fact very recently a theory has been proposed based on the idea that the free electrons in a metal lattice are to a large extent in a peculiar state which is termed "degenerate" at ordinary temperatures. It is only the electrons which are not in the degenerate state which can partake in the equilibria discussed. On raising the temperature the degeneration of the electrons is decreased, and the increase in free electrons and possibly many of the unexplained facts of metallic conduction will be much clarified by this theory which has been proposed by A. Sommerfeld, (see page 329 Chapter XXVII).

As a result of these phenomena we must now extend our Joule's

law of heating to take account of the Peltier effect. The heat liberated in a circuit containing a junction if we include this effect is represented by the equation $H = 0.24(i^2Rt \pm \pi it)$. π is a constant called the Peltier coefficient. The equation gives the heat evolved for a quantity of electricity (it) when multiplied by 0.24 at each junction, the positive and negative signs applying to junctions where heating and cooling occur respectively.

The laws of thermal E.M.F.'s are as follow:

(1) The thermal E.M.F. for any temperature of the metal junctions is always the same, provided the metals used are sufficiently pure and in the same crystalline state.

(2) For small temperature ranges, the thermal E.M.F. is proportional to the difference in temperature. This law leads to the thermocouple, which is used in temperature measurement.

(3) The algebraic sum of all the thermal E.M.F.'s in a circuit will give the total E.M.F. if the thermal E.M.F.'s are added with due regard to sign.

64. THERMOELECTRIC POWER AND CALCULATION OF THERMAL E.M.F.'S

Investigation has shown that the metals may be arranged in a series giving their relative *thermoelectric power. This term, thermoelectric power, means the electromotive force per degree Centigrade difference in temperature for a thermocouple made of the given metal and lead.* Thus the thermoelectric powers of the metals are defined in comparison with an *arbitrarily chosen standard substance—lead.* The reason why lead is used will become obvious later. On the basis of this definition we have the following list of thermoelectric powers. It must be emphasized, however, that these refer to specific samples of *supposedly pure* metals. Small traces of impurities may alter the thermoelectric powers so that the table may be completely different.

Although the values of the powers are given as independent of temperature, the *power actually does vary with temperature.* Thus for iron, the thermoelectric power is given by the expression $+ 17.5 - 0.048t$, and for copper it is given by the expression $+ 1.34 + 0.01t$. This is due to the fact that all metals but lead show an electromotive force when a temperature gradient exists in a wire. The history of the discovery of this fact is quite interesting. It is one of the few cases where thermodynamic reasoning has led to a new discovery of fact, though it serves a most valuable function in the systematization of physics.

When the effects we are discussing were first observed and measured, Lord Kelvin made a thermodynamic study of the circuit.

THERMOELECTROMOTIVE POWERS FOR THE ELEMENTS AT 20° C AND AS A FUNCTION OF TEMPERATURE t IN DEGREES CENTIGRADE.

(The standard metal in these couples is lead and the power is in volts per degree.)

Element	Power at 20° C	Power at t° C
Al	-0.68×10^{-6}	$(-0.76 + 0.0039t) \times 10^{-6}$
Au.....	$+3.0 \times 10^{-6}$	$(+2.8 + 0.0101t) \times 10^{-6}$
Sb.... .	$+6.0 \times 10^{-6}$	
Bi—(commercial).	-97.0×10^{-6}	
Bi—(pure).	-89.0×10^{-6}	
Cd. . ..	$+3.48 \times 10^{-6}$	$(+2.63 + 0.0424t) \times 10^{-6}$
Cu	$+1.52 \times 10^{-6}$	$(+1.34 + 0.0094t) \times 10^{-6}$
Fe. . .	$+16.2 \times 10^{-6}$	$(+17.15 - 0.0482t) \times 10^{-6}$
Ni—(−18° C. to 175° C).	-22.8×10^{-6}	$(-21.8 - 0.0506t) \times 10^{-6}$
Pt—(hard) .	$+2.42 \times 10^{-6}$	$(+2.57 - 0.007t) \times 10^{-6}$
Pt—(malleable)	-0.818×10^{-6}	$(-0.60 - 0.0109t) \times 10^{-6}$
Pt—Ir alloy 85% Pt − 15% Ir	$+8.03 \times 10^{-6}$	$(+7.90 + 0.0062t) \times 10^{-6}$
Pb...... .. .	0	0
Se	$+807.0 \times 10^{-6}$	
Ag. ..	$+2.41 \times 10^{-6}$	$(+2.12 + 0.147t) \times 10^{-6}$
Steel. . .	$+10.62 \times 10^{-6}$	$(+11.27 - 0.0325t) \times 10^{-6}$
Sn...	-0.33×10^{-6}	$(-0.43 + 0.0055t) \times 10^{-6}$
Zn... .	$+2.79 \times 10^{-6}$	$(+2.32 + 0.0238t) \times 10^{-6}$

From his study of the effect he found that the temperature coefficient of the thermoelectric force could not be accounted for thermodynamically unless one assumed that there were other sources of electromotive force in the circuit which were a function of the temperature. From this Lord Kelvin predicted that if one took a piece of copper wire and established a temperature gradient down its length one would observe, on passing a current through it, a heating on one edge of the temperature gradient and a cooling on the other. Following the prediction of Lord Kelvin the existence of this effect was actually verified, and it has been called after him, the Thomson effect.

The fact that the thermoelectric power and hence the thermal E.M.F. varies with the temperature, and that the variation for each metal can be plotted as a function of the temperature, has led to a very ingenious device for computing the thermal E.M.F. given by any couple as a function of the temperature.* Before proceeding

* In fact since the E.M.F. E varies with t the only way we can handle the general problem is to choose an interval dt (in this case 1° C) so small that $E = P$ is sensibly constant and take the integral of $dE, \int dE = \int P dt = E$ over the temperature range considered.

with the diagram we might mention that the reason that the thermoelectric power was chosen in terms of lead was that lead is the one element which apparently does not show any variation in its thermoelectric effect with temperature, i.e., the so-called Thomson effect which Lord Kelvin discovered is zero in the case of lead.

In Fig. 52, we have plotted the *thermoelectromotive power* as ordinates against the temperature in degrees C as abscissae. In such a diagram lead is taken as zero and is parallel to the axis of abscissae.

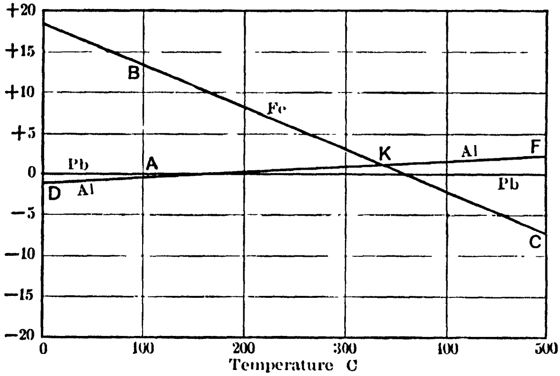


Fig. 52.—Thermoelectromotive Power against Temperature in Degrees C. (Thermoelectromotive Power to be Multiplied by 10^{-6}).

From the equation given for iron, the iron will give a straight line as depicted in the diagram. Aluminum will give another straight line which starts with a nearly zero value at 0° and slowly increases to 500°. The two curves cross each other in a point K. The problem now arises to determine the *thermal electromotive force* given by an aluminum-iron thermocouple as a function of the temperature from such a plot.

Assume that one of the junctions of the couple is at 0° C, and that the other junction is at a temperature for instance of 100° C. Were we dealing with a lead-iron thermocouple and were the thermoelectric power of the iron relative to the lead constant it would be represented by a straight line parallel to the axis of abscissae. If then, the *thermoelectric power were* 15×10^{-6} per degree at 100°, the electromotive force for one junction at 0° and the other at 100 would be $100° \times 15 \times 10^{-6}$, that is, it would be equal to the area between the lead and the iron curve lying between 0° and 100°. Since the iron has a positive *thermoelectric power* that is *not constant with temperature*, but which decreases as temperature increases, the electromotive force given with a difference in temperature between the two junctions will be given by the area between the iron and lead curves between the temperatures of the two junctions. In other words, the electromotive force in the case of any junction of two metals will be the area included between the two curves and the temperatures of the two junctions of the thermocouple.

Returning now to the iron-aluminum couple, we see that the *thermal electromotive force* observed with one junction at 100° and the other one at 0° will be given by the area included in the quadrilateral *DAB* 17. If the *thermal electromotive force* at 330° were required it would be the area included in the triangle *DK* 17. It is observed that since the aluminum is more negative than the iron in this region, the E.M.F. will have the aluminum negative and the iron positive, and since in the triangle *DK* 17 the area is the maximum possible area with the iron positive and the aluminum negative relative to it, *the point K should mark the maximum of the thermal E.M.F. curve as a function of temperature.* The 0° junction *has iron positive to aluminum by* 17. At 500°, the iron is negative to aluminum by about 8 units. Thus polarity in the two junctions is now reversed and when we add these polarities to obtain the E.M.F. we see at once that the E.M.F. will be less when the temperature difference is 500° and 0° than it was between 300° and 0°. Thus, in order to get the thermal E.M.F. from the diagram at 500° we must take the difference in the areas of the triangles 17 *KD* and *CKF*. If we compute the calculated thermal E.M.F. for each temperature by taking the areas of the triangles or quadrilaterals formed by the two lines and the temperature chosen, one would find that the thermal E.M.F. would lie on a curve which is a parabola. The vertex of this parabola would be at a temperature corresponding to the point *K* in the thermoelectromotive power diagram. A parabola such as would be observed is shown in Fig. 53.

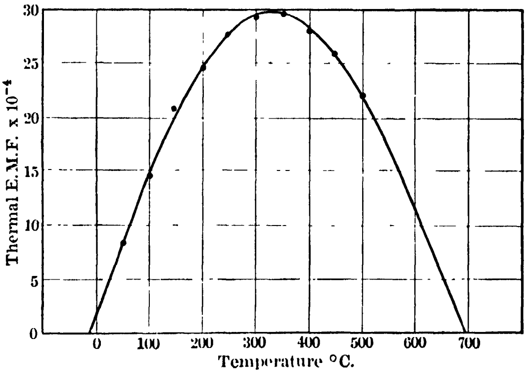


FIG. 53.—Thermoelectromotive Force of an Al – Fe Thermocouple as a Function of Temperature with Cold Junction at – 15° C.

It is obvious in the case of the iron and copper that if one went to still higher temperatures than outlined, eventually the area of the right-hand triangle would become equal to or greater than the area of the left-hand triangle. This merely means that the thermal E.M.F. would not only drop to 0, but would become the reverse of what had been observed for smaller temperature differences.

The thermal E.M.F. which we have calculated by this ingenious

device of Lord Kelvin may also be calculated directly from the formula which expresses the same relation. The equation can be easily derived from a generalization of the geometry of the system. The E.M.F. is given by

$$E = P(t_H - t_c)\left(1 - \frac{\frac{1}{2}(t_H - t)}{n - t_c}\right).$$

In this equation P is the thermoelectric power in volts per degree C at t_c for the couple, which is to be taken from the table for the two elements considered. n is the temperature of the neutral point K, and t_H and t_c are the temperatures of the hot and cold junctions respectively. By the neutral point n we mean the temperature at which the thermal E.M.F. reaches its maximum and begins to go down. It is the point at which the thermoelectromotive power of the metals becomes equal.

65. USES OF THERMOCOUPLES

Outside of its interest from a theoretical point of view, the thermoelectric effect is of great use to us in practical applications for temperature measurement. Practically all of our temperature measurements in the regions between 300° and 1200° C are made today by the use of the thermocouple. The thermocouple is nothing but a thermal junction of two dissimilar metals, one junction of which is kept at a constant temperature, generally that of melting ice, the other junction being at the place where temperature is required. The advantages of thermocouples are that they possess a small heat capacity, that they have an almost instantaneous response to the temperature because of their high conductivity and low heat capacity, that they can be inserted into positions inaccessible to ordinary thermometers, that if they are carefully calibrated they will give temperatures as accurately as we can measure the E.M.F., and finally that they can be made of metals which are sufficiently temperature resistant so that temperatures which one could not study with ordinary substances such as glass can be easily studied. For low temperature work, combinations of copper and certain alloys are used such as constantan or as iron and nickel couples. For somewhat higher temperatures two alloys known as chromal and alumal are used. These are nickel-iron alloys containing chromium and aluminum. For very high temperatures thermocouples are made of platinum. Generally the couples consist of platinum and platinum iridium or platinum rhodium alloys. Such couples must be standardized by

measuring the thermal E.M.F. at the melting points of metals whose melting points are accurately known.

Another use of the thermocouple is in the form of the thermopile. This merely consists of a group of thermocouples in series, the cold junctions all being fastened to a heavy cold block. The junctions to be warmed (i.e., the alternate junctions) are blackened and placed close together. This arrangement adds the E.M.F. of the separate couples and multiplies the effect of one couple. Thermopiles have been used together with a sensitive galvanometer mentioned in a previous chapter for measuring the radiation from stars. These piles are made of very fine metal strips and have a very small heat capacity. Another form of instrument of this type is the radio-micrometer due to Boys. This consists of a loop composed half of a fine nickel wire and half of a fine copper wire, the cold junction being fastened to a light metal bar suspended by a fiber as shown in Fig. 54. The

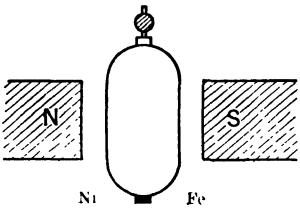


FIG. 54.—Boy's Radio-micrometer.

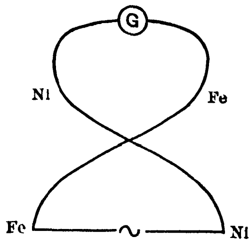


FIG. 55.—The Thermal Cross.

hot junctions consists of the other ends of the wires welded together and coated with soot. This loop of the two dissimilar wires is suspended in a magnetic field. When radiation falls on the blackened junction a current flows through the system and a deflection results. With this instrument Boys detected the radiant heat of a candle at a distance of two miles.

Finally, there is the ingenious device known as the thermal cross which measures small electrical currents, illustrated in Fig. 55. Two very fine wires, say of nickel and iron, are bent in the form of a cross, the center of the cross having the nickel wire either soldered to, welded to, or twisted around the iron wire. These wires are very fine. If a current passes through these wires, the heat at the junction produces a thermal E.M.F. between nickel and iron which can be read by a galvanometer. As the heat can be produced by high frequency alternating currents which can not be read by a galvanometer, the method is used in the measurement of the very feeble alternating currents which are frequently encountered in radio.

CHAPTER XIV

STATIC ELECTRICITY I—CHARGES AND FIELDS

66. ELECTRIFICATION AND ITS DETECTION

As was stated in the introduction, electrification was known to the Greeks. It was later studied by Gilbert, physician to Queen Elizabeth. Gilbert found that besides amber and jet many other non-conductors when rubbed with other non-conductors were able to pick up light objects. Electrification therefore manifests itself by attractions, and it is only by a manifestation of the attractive and repulsive forces that we know of the existence of electricity. It was later found that the forces between electrified particles are not only those of attraction but also that they are forces of repulsion. Shortly after Gilbert's time two kinds of electricity were discovered and named. They were typified by the electrification produced when a resinous substance like sealing-wax or hard rubber was rubbed with wool or cat's fur, and when substances like glass were rubbed with silk. The electrification obtained from resinous materials was called *resinous* electricity. The electrification obtained from glass-like substances was called *vitreous* electricity. Investigation showed that two bodies charged with the same kind of electrification repelled each other, while two bodies charged with the two different sorts of electrification attracted each other. The repulsive force between two charges was made use of in the so-called electroscope which was early used to measure and study charges. In Fig. 56, B is a metal box with glass windows on two sides. W is a wire passing through an insulated plug A consisting either of amber or sulfur. L is a small piece of gold leaf fastened to W at one point. When vitreous or resinous electrification is put on W the like charge which reaches L causes L to be repelled and to give a deflection which will later be seen to measure the electric potential. It was found that, if we had an uncharged insulated

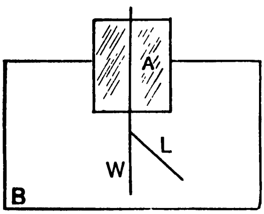


FIG. 56.—Gold Leaf Electroscope.

body and connected it by a metal wire to a charged body, the charge would travel along the wire to the uncharged body. This was called *electrical conduction.* The ability of electricity to move along conductors differentiates it from magnetism in which the two magnetic entities appear in equal amounts and are fixed in the material in which they occur. The conductivity of electricity enables us to measure experimentally the units of electricity and compare them much more accurately and easily than is the case for magnetism.

67. ELECTROSTATIC INDUCTION

If a charged body be brought near a neutral uncharged one, the uncharged body exhibits charges at its ends, as shown in Fig. 57*A*.

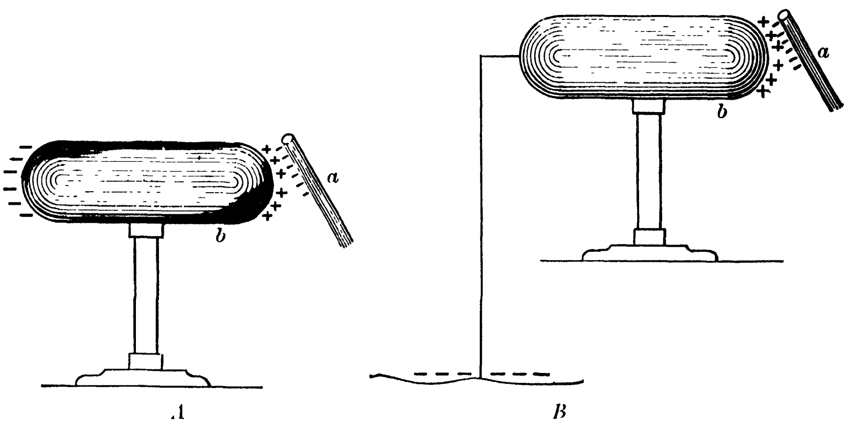


FIG. 57.—Charging by Induction.

On the end nearest the charged body it has electrification of the opposite type; on the farther end it has electrification of the same type as indicated in the diagram of Fig. 57*A*. *This phenomenon is known as induction.* If the charged body be brought up to the insulated, uncharged body, then if this be touched by a conductor connected to earth, its contact next being broken, and the charged body removed from the neighborhood of the uncharged body, the uncharged body will now be observed to have a charge. The mechanism of charging by induction is simple and is easily understood by regarding the two diagrams of Fig. 57. In diagram *A*, the resinous electricity of the charged body *a* attracted a charge of the opposite sign on the near side of the insulated conductor *b*. The charge of like sign attempted to get as far away from the charged body as possible. Making contact between the uncharged body and the ground gave the charge of

the same sign as the charged body the opportunity of getting further away, as shown in diagram *B* of Fig. 57. On breaking the contact while the charged body *a* was near the insulated body the vitreous electrification could not return to *b*. Therefore, on removing the charged body *a*, the electrification of opposite sign to the charged body on the insulated body *b* remained and spread over the surface of the insulated body. Thus, we can charge one body with the electrification of opposite sign by bringing a charged body near it, touching it, breaking contact and removing the charged body. *This is known as charging by induction.* It has very important consequences. One of them is that if the quantities of electrification produced be measured, the electrification of both kinds produced by induction in a body appears in equal amounts. That this must be so follows from the fact that before bringing the charged body near the uncharged body, the body *b* was as a whole neutral.

68. ONE AND TWO FLUID THEORIES OF ELECTRICITY

The phenomenon described above together with the conduction of electricity led Franklin to postulate a theory of electricity. He assumed that there was one electrical fluid * which he called positive electricity. It was identified with the vitreous electrification. Then there was another type of electrification which was the result of the removal of the positive electrical fluid from neutral bodies. Electricity was looked on as a fluid because it could move through a conductor. In contradistinction to the positive electrification the residual electricity was called negative. At about the same period another theory of electricity had been put forward which was that both positive and negative electricities were separate fluids, then a neutral body contained equal amounts of the positive and negative fluids. *Today*

* It is interesting to note that in the period during which these researches developed, beginning with 1750 and extending well through the succeeding 150 years, there was a peculiar type of view point in the study of matter. This view point came from the attempt to explain the universe in terms of a mechanical system based on Newton's laws. With this attempted "explanation" there developed, in order to solve the problems of motion involved, the mathematical methods of the calculus. Because of the ease with which problems could be handled by the calculus, which deals with continuously varying functions, the attempts to describe nature were all made in terms of concepts based on the continuous. Thus heat was supposed to be a weightless fluid, as was the case with electricity. In fact it has been only since 1890 that we have begun to realize the great importance of the discontinuous in physics and the era which has given us the atom, the electron, the proton, and the quantum of energy is making us search hard for an adequate mathematical and physical system of representation of a world which we find to be made up of particles instead of continuous fluids.

we know that in the sense of a single fluid the Franklin theory was correct as regards metals. One of the types of electricity, the negative electricity, is today known to consist of minute particles (not a continuous fluid), which are called electrons. Because of their small size and mass they are free to move. *They then constitute the "fluid" postulated by Franklin* but are negative in sign, and consist of particles. The atoms of the matter, when the negative electrons are removed, are positively charged. They constitute the residual "negative" charge of Franklin. *Where the atoms and molecules are free to move as in liquids, we have electricity moving according to the two fluid theory.* Hence, the bitter controversy which raged over this question was entirely futile and unnecessary, for both views have been shown to be partly right and partly wrong.

69. CONSEQUENCES OF INDUCTION—THE ICE-PAIL EXPERIMENT

The phenomenon of induction leads to several important applications. The first one is the phenomenon of electrical screening. Suppose one place inside of a box a charged metal body, then on introducing the charged metal body the electricity of opposite sign will flow to the inside of the box tending to neutralize the charge placed inside. The electrification of the same sign as the charged body will flow to the outside of the box, and if the box be attached to earth will flow away. Thus, outside of the box there will be no manifestation of electrification, the electrification inside of the box neutralizing the field of the charge.

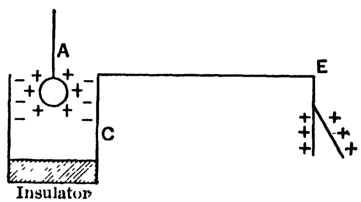


FIG. 58.—Faraday's Ice-Pail Experiment.

This *screening effect*, so called, is proven by Faraday's famous ice-pail experiment, shown in Fig. 58. In this experiment, the charged body A with positive electricity is brought inside of a metal cup C which is connected to the electroscope E. As a result of the positive charge, negative electrification will be drawn to the part of the cup nearest A. The positive electricity will then remain on E causing the leaves of electroscope to diverge. If now the *positively charged body A be touched to the inside wall of the cup C* and then removed, *no change will be observed in the leaves of the electroscope.* They will stay separated exactly as far as before. This means that *the induced positive charge on the electroscope is exactly the same as the charge that was communicated to the electroscope by A.* Thus one sees that the charge on A was completely neutralized by the negative charge which resided

in C, the equivalent of this in positive charge being given to E. The consequences of the ice-pail experiment are greater than they would seem at first sight, for besides leading to the phenomenon of screening, *they lead to a proof of the law of force between electrical charges* of very great precision. Of this, more will be said later.

70. USES OF INDUCTION

Induction can be made use of for producing electrification. In its simplest form induction is achieved by the instrument known as the *electrophorus*. This consists of an insulating plate, for instance, of hard rubber which can be charged by friction. An insulated metal disc is brought as near as possible to the charged plate without making contact except at one or two points. The metal disc is then touched for an instant giving the electricity of the same sign as the charged ebonite plate a chance to flow off. The finger is then removed and when the metal disc is removed it will be found to be highly electrified. This electricity can be communicated to any system, the plate thus discharged, and being again brought near the charged ebonite can be charged again. Thus the process of charging and discharging can be carried on indefinitely with the same quantity of electricity on the ebonite. *The energy for producing this electrification is furnished by the work of bringing the metal disc into the neighborhood of the ebonite plate and lifting it away against the attractive forces.* Thus one does not get electric energy for nothing. In fact, the energy obtained this way under ideal conditions would be equal to the work done. Mechanical devices to perform this operation continuously, which store the electricity obtained in proper containers, have been invented. They are called electrostatic machines, and served in the early days as the only source of electrification. They were also used in the early days to get high voltages for x-ray work. Today, we have far more constant and efficient sources.

71. QUANTITATIVE TREATMENT—COULOMB'S LAW

Having discussed the qualitative facts of static electricity it now becomes necessary to take up the quantitative facts. As the phenomenon of electrification is characterized by attractions and repulsions we regard electrification as being a manifestation of the action of forces. *These forces differ from forces such as gravitation in that they are produced on bodies by different means. Forces of the type above described, that are produced by frictional processes, by processes going on in chemical solutions, or by the action of magnetic fields on conductors,*

we define as electrical forces. The only way in which we can measure electrification is in terms of such forces. It is therefore essential that we know the law of force acting between electrical charges and use this law to define our unit of electrification. *Coulomb,* 1785, *using the torsion balance, was the one to formulate the laws of electrostatic attraction and repulsion.* Cavendish had also deduced this law at an earlier date (1762). As he did not publish his results, the credit has been given to Coulomb. Coulomb found that the force was proportional to the product of the charges on the two bodies and that it was inversely proportional to the square of the distance. It was later observed that the force was inversely proportional to a constant of the medium in which the charges were immersed. This constant, designated by the letter D, is called the *dielectric constant.* More will be said of this in a later chapter. Symbolically we write this law

$$f = \frac{qq'}{Dr^2}.$$

In the above equation q and q' are the two charges and r is the distance between them, while f is the force.

That the force varies inversely as the square of the distance is further proven by the ice-pail experiment on the basis of the following reasoning. Assume that we have two concentric spheres as shown in Fig. 59, where A is insulated from B, which surrounds it except for a minute hole through which a wire W can make contact between B and A. If now B be charged and touched to A no decrease in the charge on B will be noted and no charge will appear on A. This result is closely similar to the one obtained in the ice-pail experiment, when the sphere produced no change in charge on being touched to the walls, but is more likely to be accurately verifiable than the cruder demonstration experiment.

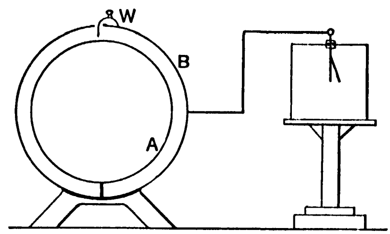


FIG. 59.—Modified Ice-Pail Experiment.

The experiment must be interpreted as meaning that there is *no electrical force on electrical charges between the outer and inner* sphere; that is, a charge on the outer sphere exerts no resultant force on the inner sphere or on charges between the inner and outer sphere, as no electricity flows between A and B in either sense when connected together. This result is perfectly general for all spheres A and B so that it may be generalized to include any point within B. We can

now idealize the situation by taking a point P at any place between A and B as in Fig. 60 and considering it alone inside the sphere B. We know that there was no resultant force on any charge at P due to the charge on B, as there was no change in the state of B on connecting B and any point P electrically. Draw through P a plane indicated in cross section by the line COD, which passes through O, the center of B. Then pass a plane EPF perpendicular to COD through P. Two cones may then be passed through P in opposite senses and of equal solid angle $d\omega$. These cones are represented in section by the two elongated triangles Px_1ds_1 and $Pxds$ intersecting the spherical surface in the small surface elements represented in section by the lines ds_1 and ds. It is seen by joining O, the center of B, to the two ends of one of the sides of the cone that the intercepted surface elements ds and ds_1 make angles $\alpha = \alpha_1$ with the perpendicular lines H and H_1, representing the right sections of the two cones. The areas of the right sections represented by H_1 and H are from geometry $r_1^2d\omega$ and $r^2d\omega$, where $d\omega$ is the solid angle of the cone and r_1 and r are the distances of P from the right sections H_1 and H at one point of each.

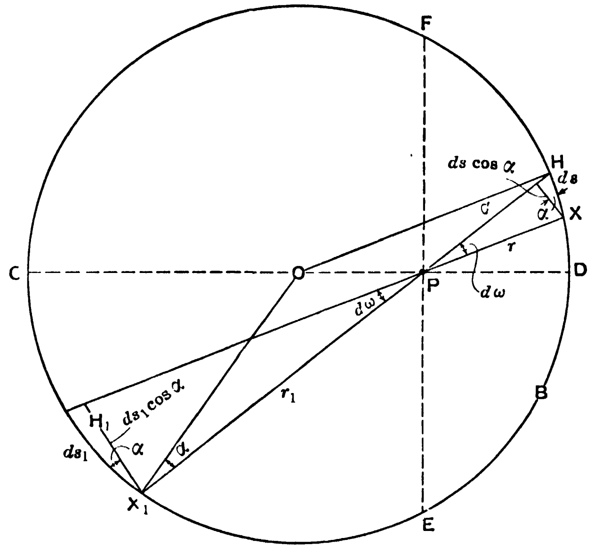


FIG. 60.—Proof of Inverse Square Law from Ice-Pail Experiment.

It then follows at once that $ds = \dfrac{r^2d\omega}{\cos\alpha}$ and $ds_1 = \dfrac{r_1{}^2d\omega}{\cos\alpha}$. Now if the charge had been uniformly distributed over the sphere B as we have assumed then let the surface density (number of charges per cm²) be σ charges per cm², and the charges on ds_1 and ds will be σds_1 and σds, or $q_1 = \dfrac{r_1{}^2\sigma d\omega}{\cos\alpha}$ and $q = \dfrac{r^2\sigma d\omega}{\cos\alpha}$. Now if a charge q_2 were placed at P it would be attracted, if Coulomb's law is correct, by a force $f_1 = \dfrac{q_1q_2}{r_1{}^2}$

towards ds_1 and by a force $f = \frac{qq_2}{r^2}$ by ds, or in one sense by a force $f_1 = \frac{q_2 r_1^2 \sigma d\omega}{r_1^2 \cos\alpha}$ and in the other sense by a force $f = \frac{q_2 r^2 \sigma d\omega}{r^2 \cos\alpha}$. It is obvious that these two forces are on this assumption independent of the values of r_1 and r, and are equal. Therefore, the resultant force on a charge at P, on the Coulomb law which say that f is proportional to $\frac{1}{r^2}$ is 0, a fact that is borne out by observation. Since the cones were chosen completely at random as was the point P this reasoning holds equally well for any other elements ds and ds_1 chosen lying on two sides of the line EPF and for any point within the sphere B. The conclusion which we draw, therefore, is that if Coulomb's law holds there will be no force on a quantity P placed anywhere within sphere B acting to move it towards or away from O the center of B. Thus a charge on a wire between the spheres A and B would neither move towards A nor B as was observed to be the case.

Had, however, the law of force been $f = \frac{q_2 q_1}{r^n}$, where n could have a value different from 2, the reasoning would have lead us to the conclusion that

$$f_1 = \frac{q_2 r_1^2 \sigma d\omega}{r_1^n} \text{ and } f = \frac{q_2 r^2 \sigma d\omega}{r^n}, \text{ or } f_1 = q_2 \sigma d\omega\left(\frac{1}{r_1^{n-2}}\right) \text{ and } f = q_2 \sigma d\omega\left(\frac{1}{r^{n-2}}\right).$$

If then n were greater than 2, f_1 would have been less than f since r_1^{n-2} would have been greater than r^{n-2}, as r_1 is greater than r for all cones through P for all points P. Hence if n is greater than 2 the positive charge on B would move towards A, or there would have been a resultant force towards the center O. If n is less than 2, r_1^{n-2} and r^{n-2} have negative exponents and f_1 will be greater than f. Thus in this case the positive charge would have moved from A to B, i.e., away from the center O. As a result of measurements which are very sensitive to small changes in potential for A while the value of σ on B is very great we can show that n has the value 2 to better than one part in many millions.

The essential notion in this proof is simple in that if we can choose conditions for test which are simple, i.e., concentric spheres where calculation is very simple, then as the area of the surfaces intercepted on the sphere by the cones drawn through any point P and hence the charge on these surfaces are in direct proportion to the squares of the distances from the point P, there must be a resultant force on P

towards or away from the center of the sphere unless the law of force between the surface charges and the point varies inversely as the square of the distance. *This reasoning shows that the only law of force compatible with the fact that a charge inside of a conducting body exerts no force, is the inverse square law of force.* The experimental proof of the extent to which this law holds was due originally to Bertrand, Maxwell, and Faraday. The latter went inside a conductor charged to potentials of hundreds or thousands of volts with the most sensitive electrometer available and discovered no field inside the conductor. The above mentioned experiment establishes the inverse square law very accurately for fields of considerable magnitude. The range of distance between bodies over which the inverse square law has been observed to hold is enormous. It holds, as far as we can tell for distances of several meters, and the recent experiments of Rutherford on the deflection of rapidly moving positively charged atoms of helium by the positive nuclei of the elements has proven that this law holds true within one percent down to distances of 10^{-12} cm. Beyond this, the law ceases to hold and these dimensions are supposed to indicate the order of magnitude of the diameters of the units of electricity, the electrons and nuclei.

Returning now to the expression for the force given above, if we let q equal q', a result easily achieved, for instance, by touching two charged spheres of equal size together, we have

$$q = \sqrt{fr^2},$$

when D * equals 1. (Actually D is unity for empty space only. In air, however, it is so near to 1 (1.0006) that we can call it unity.) We now say that *q equals 1 electrostatic unit of electricity when r equals 1 cm and f equals 1 dyne. This gives an arbitrary definition of unit electrostatic-quantity in terms of the force produced at the unit distance, as is shown by the value of the charge on the true unit of electricity whose value is 4.77×10^{-10} of these electrostatic units.*

In words, we can define the unit quantity of electricity on the electrostatic system as *the quantity of electricity which repels an exactly equal quantity of electricity at a distance of 1 cm with the force of 1 dyne (in vacuo).*

It is seen that this unit is *small*, for a dyne is a weak force. The electrostatic unit of quantity might then be expected to be less than the electromagnetic unit of quantity, and in fact 1 *electromagnetic unit of quantity* $= 3 \times 10^{10}$ *electrostatic units of quantity*, and *one coulomb equals* 3×10^9 *electrostatic units of quantity.*

* The question of the dimensions of D is treated on page 69.

72. DEFINITION OF CURRENT IN THE ELECTROSTATIC SYSTEM

We have now defined the quantity of electricity in terms of the electrostatic system of units. This enables us at once to use this concept to define current in terms of the electrostatic system; that is, to define current in terms of a new set of phenomena. In the electromagnetic system, we defined current by the magnetic field produced and defined quantity of electricity as current times the time. Now, we are able to define *quantity* in the electrostatic system with great ease. The magnetic effects of the usual currents obtained by static means are, however, far too small to detect and measure accurately. That a current from an electrostatic generator causes a magnetic field can be easily proven experimentally. Since the unit of quantity in the electrostatic system is very small, the unit of current (quantity per unit time) is also small so that this unit is not convenient as a means of evaluating currents that are ordinarily measured in the laboratory. It is accordingly rarely encountered except where very small currents of electricity are involved. *Current, on the electrostatic system,* as indicated above, *is merely defined as the quantity of electricity which has been transported divided by the time taken to transport it.* Symbolically, it is

$$i_a = \frac{q}{t}.$$

If q is one electrostatic unit, and t is one second, i_a is one absolute unit of electrostatic current. It is related to the electromagnetic unit of current in that it is

$$\frac{1}{3 \times 10^{10}}$$

electromagnetic units, or 1 E.M.U. of current = 3×10^{10} E.S.U. of current. The significance of the value of this ratio has been shown in Chapter III. The ampere is accordingly 3×10^9 E.S.U. of current.

73. DEFINITION OF ELECTRIC FIELD STRENGTH

It is obvious, therefore, that with a charged body, there is at any point in space surrounding it a force manifested on another charged body placed at that point. We can thus say that the charged body has a field of force about it. Again, in electricity as in magnetism, the intensity of this field of force at a point is represented by *the force in dynes which would be exerted on a unit positive charge placed at that point in the field.* Thus electrical field strength F at a point is defined

as the force per unit charge placed at that point.* Such a force has magnitude and direction. It is therefore *a vector quantity.* We can consequently determine the force at any point in the field due to a number of charges by adding *vectorially* the separate forces produced at that point due to the separate charges. Again, as in magnetism, we may map the field of force and conventionally we *represent the unit field* as *a field having one line of force per cm² of area taken perpendicular to the direction of the field.* That is, a field of force which would act on a unit positive charge at a point with a force of 1 dyne will be represented by 1 line of force per cm^2.

One cm from an isolated charge q, the force on a unit charge is q dynes, that is, the field at 1 cm distance is q dynes in strength. Since a spherical surface surrounding the charge at 1 cm is $4\pi cm^2$ there will be $4\pi q$ lines of force emerging through the surface; that is, *a quantity of electricity q sends out $4\pi q$ lines of force.*

* In the case of the field due to an isolated charge q, Coulomb's law says that the force f on a charge q' distant r cm is $f = \frac{qq'}{r^2}$. The field F is then given by $F = \frac{f}{q'} = \frac{q}{r^2}$.

CHAPTER XV

STATIC ELECTRICITY II—POTENTIAL

74. DISCUSSION OF POTENTIAL IN THE ELECTROSTATIC SYSTEM

WE now come to a new unit in the electrostatic system. We showed in the last chapter that at any point in the neighborhood of an isolated charge q there is a force on another charge q. Accordingly about a positively charged body q at Q there is a field of electrical force surrounding the charge q. Suppose we have a quantity of electricity q' at a point A in the field of q and assume that q' is positive, as shown in Fig. 61. Let us move $+ q'$ from A to another point B nearer Q. Since the force is one of repulsion, we must do work on the charge $+ q'$ to move it. Now, since the strength of the field at a point distant r from a charge q is

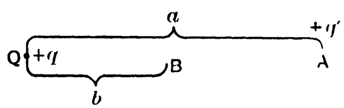


FIG. 61.

$$F = \frac{q}{r^2},$$

the work done in moving a quantity q' a distance dr from A toward B is given by

$$dw = \text{force} \times \text{distance} = Fq'dr = \frac{qq'}{r^2}dr.$$

Thus the total work for all elements of path dr in going from A to B is given by

$$W_{ABq'} = \int_a^b dw = qq'\left(\frac{1}{b} - \frac{1}{a}\right).$$

If q' were equal to unity, this would be the work done to bring a unit positive charge from A to B, or the work done to move unit positive charge from one point to another in the field. But the work to move unit charge from one point to another in the field was called the potential difference between the two points when we were discussing the electromagnetic system. Therefore, the potential difference in the

electrostatic system between the points A and B, distant from a charge q by the distances a and b respectively, is

$$PD_{AB} = q\left(\frac{1}{b} - \frac{1}{a}\right).$$

It is to be noted that this is the *potential difference* between the two points. If a had been infinity, $1/a$ would have been 0 and the potential at B would have been q/b. This latter potential is called the *absolute* potential of the point, for it is the work done to carry a charge up to the point in question from infinity. In general, we measure *potential difference and not absolute potential*, as we have no point of reference other than that of the earth, and we do not know the *absolute potential* of the earth. In fact, we arbitrarily call the potential of the earth 0.

The use of potential has certain great advantages. In general, potential P is given by

$$P = \int_{\infty}^{r} F dr,$$

where F is field strength and dr is an element of length in the direction of the field; that is to say, the absolute potential of a point is the integral of the field strength over each element dr of the path times the length of the element dr taken from infinity up to the point. Because of the relation which is quoted we can reciprocally write that the field strength $F = \dfrac{dP}{dr}$ at any point in question. Thus field strength is the derivative of the potential along the line of the acting force at the point in question, or better, it is the *rate of change of potential with the distance at the point in question.* Thus, the units used in designating electrical field can be expressed in terms of units of potential difference per unit of length. The most commonly used unit for electrical field strength is the volt per cm. Fields are sometimes also spoken of in electrostatic units of potential per cm. Now, a study of the dimensions of potential would show that it is a *scalar* quantity, that is, that it has magnitude but is in no sense dependent on direction. Force is a vector quantity. Consequently, the addition and the mathematical treatment of forces due to charges involves the complication of vector quantities. For many problems, it suffices to carry through the calculations in terms of the potentials and then from them to derive the forces by differentiation. Like forces, the potentials at any point are the sum of the separate potentials due to the separate charges acting at that point taken with due regard to

sign. It is therefore obvious that the knowledge of these relations is of considerable importance and may serve to simplify many otherwise difficult problems.

75. UNIT OF POTENTIAL DIFFERENCE DEFINED

We regard potential as being positive when it *takes* work to bring a positive charge to the point in question from infinity. The unit of potential difference is defined in terms of the work it takes to bring unit positive charge from a point A to a point B between which the potential difference is sought. If we are dealing with the electrostatic system, *the unit of potential difference in the electrostatic system is the potential difference which requires that one erg of work be done when a unit positive electrostatic charge is brought from one point to the other against the existing field.*

Now, the electrostatic unit of quantity is a small quantity of electricity in comparison to the electromagnetic unit, as we have seen. To transport such a small unit against a field and expend an erg of work in doing so requires that a fairly high potential difference exist. It is, therefore, not surprising that this *electrostatic unit of potential difference should be fairly high. The practical unit which was chosen as the volt is* $\frac{1}{300}$ *of an electrostatic unit of potential difference, for the electrostatic unit of potential difference happens to be* 3×10^{10} *electromagnetic units of potential.*

Returning now to the question of the potential we notice that the potential about an isolated charge q at a distance r is equal to q/r. It is obvious that as long as r is constant and q is constant the potential is constant. That is, the potential over the surface of a sphere of radius r drawn about q is a constant and is determined by r and q. We call such *a surface an equipotential surface.* This means a surface over which the potential is constant. If the potential is constant over a surface it means that we can move a charge from one point to another along the surface without doing any work, because potential refers to the work done in moving the charge. Furthermore, this means that the force is everywhere perpendicular to the equipotential surface. For were it parallel to the surface, or did it have a component parallel to the surface, then in moving a charge from one point to another along the surface work would be done against this component and the surface would not be equipotential.

A surface of metal, being a conducting surface, allows the charges to move freely and arrange themselves in such a manner that there is no force parallel to the surface, for as long as there is a force parallel

to the surface, the charges will move. Thus, when the charges have come to rest on the surface the surface must be equipotential. Hence the potential of an isolated conducting positive sphere of radius r, carrying a charge of $+ q$ and infinitely distant from any other charges, is q/r.

Another important fact to be taken account of is that the work done in carrying a charge between two points in an electrical field is independent of the path. If work were not independent of the path in going from one point A to another point B, then we could arrange such a cycle that in going from A to B we took a path of less work, and in going back from B to A we took a path of greater work. If one had performed such a cycle one would have gone around the path AB and gotten more work in going from B to A than one put in in going from A to B. One would thus get energy at the expense of nothing, which is contrary to the first law of thermodynamics. Another way of proving this would be to choose any particular path in going from A to B which one wished (see Fig. 62). If then one took the components of the path parallel to the field or the components of the path normal to the equipotential surfaces between the two points and added them up one would find that the sum of these normal components would be equal, independent of the path chosen. But these normal components represent the components along which work was done in going from one point to the other. Therefore, the work done in going between the two points is independent of the path.

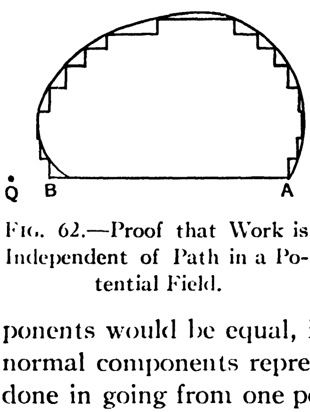


FIG. 62.—Proof that Work is Independent of Path in a Potential Field.

76. CHARGE DENSITY AND ITS CONSEQUENCES

There is one other important concept which may be used. Assume that we have an electrified surface such as the surface of a charged metal sphere. We can look at the electrification as residing on the surface and we know that this electrification consists of individual units of electricity distributed over the surface. We then speak of the *surface density of charge*, by which we mean the number of units of quantity per cm^2 of surface. The concept of surface density is sometimes useful in the treatment of problems.

An isolated sphere, for which the curvature is uniform, will have a uniform distribution of charge over its surface. In fact, the field about the spherical charged conductor is also uniform. In other bodies this will not be the case, as we shall see.

For the case of the conducting sphere of radius r, since the surface is an equipotential surface, we can draw some further important conclusions. Suppose we have σ charges per cm^2 on the surface of the sphere. The number of lines of force emerging from this surface will be $4\pi\sigma$ per cm^2. This at once enables us to calculate the field close to the surface of the conductor as $4\pi\sigma$, for there are $4\pi\sigma$ lines of force leaving the surface per square cm normally.

Since there are $4\pi\sigma$ lines emerging per cm^2 from the surface, the total number of lines of force leaving the sphere will be $4\pi\sigma(4\pi r^2)$. If now we consider that the charges are concentrated at the center of the sphere and consider the sphere removed, we would find that the spherical surface of radius r about the point where the charge is concentrated would have emerging from it a total of $4\pi(\sigma 4\pi r^2)$ lines of force and the number of lines per square cm emerging through the surface of the sphere of radius r would be exactly the same as was the case with the conducting sphere of radius r. We can accordingly in the treatment of problems consider a charged conducting sphere as being replaced by a point charge with the *total quantity of electricity on the surface of the sphere concentrated at the center of the sphere.* As can be imagined, such a treatment of spheres may simplify many calculations such as those of forces due to charged conducting spheres.

Again the simple conditions existing for the case of the surface of a charged sphere can be applied to the study of other surfaces. Any curved surface can be more or less closely approximated, at a given point, by the surface of a sphere of appropriate radius. A charged equipotential surface of any curvature can for a small fraction of its area be approximated to the surface of a conducting sphere having the same radius of curvature. Since we know the potential of the surface of a charged conducting sphere of given radius as well as the field strength close to the surface if the charge density is given, we can apply this knowledge to evaluating the potential and field strength at any point near the surface of a charged body if the radius of curvature of the approximating sphere at the point is known. Reciprocally, knowing the radius of curvature of the approximating sphere and the potential or field strength at the surface we can calculate the surface density of charge for any curved conducting surface of equal radius of curvature.

An interesting illustration of the use of the concept of surface density is furnished by an electrified metal body of a shape such that the curvature of the surface at one end is very different from the curvature of the surface at the other, as shown in Fig. 63. If we measure the surface density at the various points of the surface of

this body, we will find that at the very sharply curved portions of the surface, the density, that is the charge per cm^2, is very much higher than at the less sharply curved portions. This fact can be shown experimentally by touching a small conducting body which is insulated, a proof plane, to the different portions of the surface, and by means of this body picking up a portion of the charge at the point touched and transferring it to an electroscope. The electroscope will take on a charge proportional to the quantity of electricity picked off from the surface by the proof plane provided that the area of contact of the proof plane with the surface was the same at each point of contact. The charge taken from the pointed end of the conductor will be observed to be far greater than the charge taken from the blunt end. The explanation of this phenomenon is as follows. Since the curved portion of the surface can be approximated to that of a sphere of appropriate radius, the potential of a curved conducting surface is equal to the quantity of electricity on it divided by the radius of curvature of the approximating sphere,

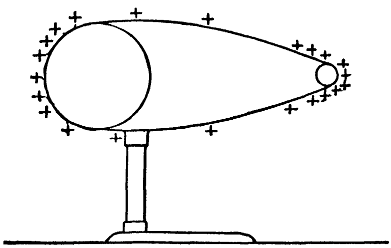


FIG. 63.—Surface Distribution of Charge on an Irregular Conducting Surface.

$$V = q/r.$$

Now, the potential V over the *surface of a conductor*, no matter what its curvature, or whether the curvature be constant, is a constant. Consequently, in order that V be maintained constant over the surface, q must vary as r varies. That is, q must be proportional to r. But the charge density is by definition q/A, where A is the area; that is, the charge density will be proportional to $1/r^2$, for a spherical surface. q is, however, proportional to r; thus q/A is proportional to r/r^2 or $1/r$. The surface density q/A, or charge per unit area, is inversely proportional to the radius of curvature. Accordingly, if we have a surface with a small radius of curvature, that is a pointed surface, the charge density will be very high, while for a blunt surface the charge density will be very low. The high value of charge densities at points makes the fields in the neighborhood of points very intense, for $F = 4\pi\sigma$. The intensity or field F becomes so great in the neighborhood of a needle point that the very few electrons existing freely in the air (about one in 10^{18} molecules), are acted on by sufficient forces to cause them to tear air molecules apart, producing

more ions and electrons. The process by which electrons generate ions in colliding with molecules is called ionization by collision and is briefly discussed in Chapter XXVI. The electrons and positive or negative ions are either repelled or attracted in the field, depending on whether the point is positive or negative. The result is that the field repelling either negative or positive ions will set the air in motion because the ions in moving through the air take the air with them, on account of the viscous drag of the air molecules on the ions. The result is that a current of air will be set up in the intense ionization produced by a high field around a point. The currents of air so produced are sufficient to cause marked mechanical effects, such as the blowing out of a flame or the spinning of a pin-wheel.

77. THE ABSOLUTE ELECTROMETER

The definition of difference of potential given on page 97 enables us to measure the potential in absolute value in terms of electrostatic units of potential. If the ratio of the electrostatic unit and the electromagnetic unit of potential is known we need not depend for our knowledge of potential on the heating measurements used in the electromagnetic system. The instrument used for the purpose is called the *absolute electrometer*. In Fig. 64, we see two parallel circular plates. The upper one has been cut so that it is formed of two parts, an inner disc and an outer annular disc with a very minute air gap G separating the two. The inner circular disc of the upper system is fastened to the arm of a very sensitive beam balance B. If, now, the balance, the upper plate and annular ring be grounded and attached to one side of the source of potential while the other insulated plate K be attached to the other source of potential a force of attraction will exist between the two plates. It will be necessary, therefore, to change the weights on the balance pan p to counteract the attractive force.

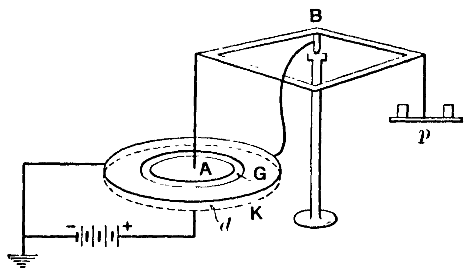


FIG. 64.—The Absolute Electrometer.

Call A the surface of the central upper plate, call d the distance between this plate and the lower one. If, now, σ is the charge density, that is, the number of units of charge per cm^2 of the plate, the total charge q on the plate of area A is $q = A\sigma$. The total number

of lines of force is equal to $4\pi q = 4\pi A\sigma$. Now, between such plates with an annular disc, which is called a *guard ring*, the field is uniform. A uniform field means that the lines of force run parallel and are perpendicular to the plates carrying the charges. The value of the field strength of such a field is the number of lines per cm², or $F = 4\pi\sigma$.* In Starling's "Electricity and Magnetism," page 132, it is proven that the force per cm² on a surface with σ charges per cm² (or a charge density σ), in a field F is given by the expression

$$f_1 = \frac{F\sigma}{2}.$$

That this must be so follows from the fact that inside the charged conductor the field is 0, i.e., a charge σ would experience no force inside the surface. On the other hand, outside of the surface the

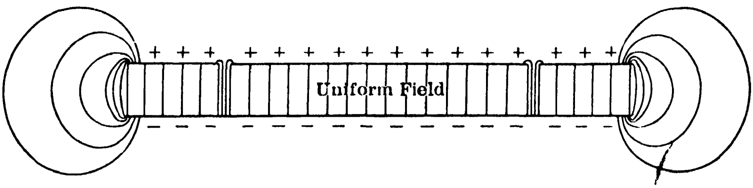


Fig. 65.—Use of a Guard Ring to Obtain a Uniform Field.

field F equals $4\pi\sigma$ very close to the surface, and a charge σ experiences a force $F\sigma$ or $4\pi\sigma^2$. A charge σ *exactly on the surface* must have a force

* To simplify our conditions where the study of quantities varying with electrical field strength is concerned we frequently resort to the use of the field produced by a pair of parallel plates whose linear dimensions are great compared to the distance between the plates. If occasion demands a small area of one of the plates, as is the case in the absolute electrometer with the plate A, the guard ring can be utilized for extending the area of the plate and maintaining the uniform field. The reason for the existence of a uniform field under these conditions lies in the fact that the greater majority of the charges are distributed on the inner surface of the plates facing each other. Under these conditions the lines of force run straight across from one plate to the other normal to the plates except near the edges, as seen in Fig. 65. In order to maintain the proper distribution of potential the charges will arrange themselves as uniformly as possible on the surfaces of the plates at some distance from the edges. The result will be that the lines of electric force run perpendicular to the plates and parallel to each other, except near the edges. The number of lines of force crossing a square centimeter of area between the plates A and K is, therefore, quite constant over the distance d between the plates, and F the field strength does not vary in going from A to K. Such uniform fields are very useful in eliminating one of the many variables in experimental investigations, namely, that of a changing field.

intermediate between 0 and $(4\pi\sigma)\ \sigma$ acting on it as a result of a field $F = 4\pi\sigma$, which is $(2\pi\sigma)\sigma$ or $\frac{F\sigma}{2}$.

Therefore

$$f_1 = 2\pi\sigma^2 = \frac{F^2}{8\pi}.$$

Again the potential difference between the plates, PD_{AK}, is

$$PD_{AK} = \int_A^K F dr = F(K - A) = Fd,$$

where d is the distance $K - A$ between the plates.

Therefore

$$F = \frac{PD_{AK}}{d}.$$

This at once gives us the total force f_A on plate A as

$$f_A = f_1 A = \frac{F^2}{8\pi} A.$$

Accordingly

$$f_A = \frac{(PD_{AK})^2}{d^2} \frac{A}{8\pi}$$

or

$$PD_{AK} = d\sqrt{\frac{f_A 8\pi}{A}}.$$

Since we can measure f_A, the electrical force, by the weights added to balance, and since we know A and d accurately, PD_{AK}, the potential difference, is given accurately in absolute *electrostatic units.*

78. GOLD LEAF ELECTROSCOPE

Frequently, for experimental work, we use a more convenient instrument known as the gold leaf electroscope for measuring the potential difference. Instruments fitted with a pivoted aluminum needle and a vertical support, similar in form to the gold leaf electroscope, only more robust, are also often used for rough potential measurements. They are known as the Braun electrostatic voltmeters. Such instruments must be *calibrated* in terms of the absolute electrometer or other potential measuring devices.

In the gold leaf electroscope, a gold leaf L, Fig. 66, is fastened to one plate A of a parallel plate condenser system. The plate A to which it is attached is insulated and connected to the source of potential to be measured. The other plate B is grounded or attached to the other side of the circuit. There is then a field F between A and B which, if it were not distorted by the gold leaf, would be given by $\frac{PD_{AB}}{d}$, where d is the distance between A and B. The charges on the leaf q' act in F to produce a force $f = Fq'$, driving the leaf from A to B. This is counteracted by the force of gravity mg acting downward to cause it to move towards A. The leaf will then take on an equilibrium position, making an angle θ with A roughly given by $\frac{f}{mg} = \tan \theta$. Thus $PD_{AB} = \frac{mgd}{q'} \tan \theta$, and one has the deflection dependent on the potential. Such an instrument, needless to say, must be calibrated to read volts.

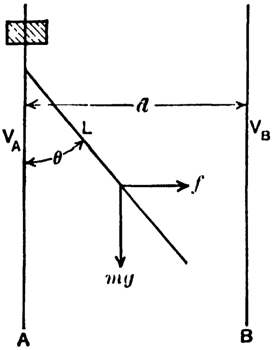


FIG. 66.—Principle of the Gold Leaf Electroscope.

79. QUADRANT ELECTROMETER

A third important instrument for measuring very small differences of potential is the quadrant electrometer. In Fig. 67, the

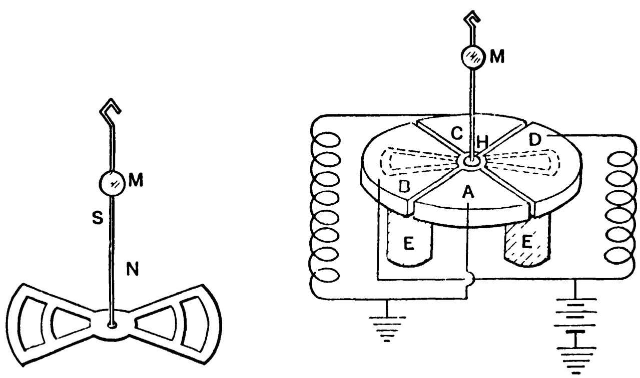


FIG. 67.—The Quadrant Electrometer.

object $ABCD$ is a cylindrical brass box with top and bottom, which has been sawed into four equal quadrants, A, B, C and D, so that each quadrant is separated from its neighbor by a small gap. A, B,

C and D are each mounted on amber insulators EE and the opposite quadrants A and C, and B and D are connected together. A very light conducting needle, the shape of which is shown in the diagram as N, is suspended inside the cylindrical box with the upright support at its center passing through the hole H in the center of the quadrants. The needle is charged to 100 volts or so by the conducting suspension, generally a silvered quartz fiber, and insulated from the quadrants. The stem S carries a light mirror M by which the deflection of the needle can be read by telescope and scale. One pair of quadrants is grounded (i.e., at zero with reference to the needle); the other pair is connected to the source of potential to be measured. When a P.D. is applied between AC and BD, the needle is attracted or repelled by the quadrant pairs in a sense depending on the sense of the P.D. and the potential on the needle. Thus the needle suffers a torque and the suspension is twisted through an angle roughly proportional to the potential difference. The whole instrument must be covered with a grounded screening case, and all leads to the sensitive quadrant must be screened with grounded shields. Some electrometers will deflect a beam of light so as to displace the spot of light reflected from the mirror on a scale at one meter's distance by 10,000 centimeters for 1 volt P.D. and have a period of only 30 seconds. The screening is necessary to avoid disturbances due to the induced static charges caused by charged objects in the room. The deflection is proportional to the potential under some conditions of operation, but the instrument needs calibration in terms of a known potential difference, which can be accomplished by the use of the potentiometer and a standard cell. For alternating potentials the needle and one pair of quadrants can be joined to one terminal, the other terminal and the second pair of quadrants being grounded. This connection is called *idiostatic* and leads to deflections proportional to the square of the potential difference.

CHAPTER XVI

STATIC ELECTRICITY III—CAPACITY

80. DEFINITION OF CAPACITY AND THE UNITS OF CAPACITY

We are now in a position to define a new *derived unit.* This derived unit becomes more prominent in the electrostatic system of units, although it occurs as well in the electromagnetic system of units, especially where currents varying with time occur. The reason for its importance in the electrostatic system is two-fold. In the first place, we deal in static electricity largely with the quantity of electricity and not with the current. Furthermore, the measurement of small quantities of electricity and of the weakest currents is achieved in a great many cases by the use of this new unit together with a measurement of the change of potential.

When we speak of the capacity of a tank for holding water we use the expression indiscriminately in two senses. In the more general use we mean by capacity the total quantity of water which the tank can hold. More properly, assuming the tank to have vertical sides, we can use the expression to denote the volume contained per unit height of water in the tank. Thus the term capacity in the latter sense could be defined as the quantity of water divided by the level to which it fills the tank, or the quantity divided by the potential. This definition makes the capacity independent of the height of the tank and defines the capacity in terms of its influence on the flow of water from the tank, i.e., the potential or head of water. With this definition the *total* capacity in the general use of the word would be limited only by the pressure which the tank could stand without bursting.

In dealing with electrical circuits and systems, the potential of a body is of first importance. It is therefore not strange that electrical capacity should be defined in analogy to the second sense above mentioned for the storage of water; that is, it is not unexpected that electrical capacity should be defined as the quantity of electricity required to be placed on a system to raise its potential by unity. The usage in this sense is also the more natural as unlike water tanks,

where special construction is needed to allow one to reach potentials or depths such as to approach breakdown, electricity can be added to a system without added construction until the breakdown point of the insulators is reached. Thus the total quantity stored in an electrical capacity is limited only by the breakdown strength of the insulating material, while in general this is not so in water tanks.

Choosing therefore the second use of the term capacity mentioned in speaking of water tanks, the analogy between the behavior of electricity and water used before is strictly maintained and applicable, and we have a definition of our new derived unit in terms of the ratio of the two chosen fundamental units, quantity and potential. *We will thus define the capacity of a system in electricity as the quantity of electricity necessary to change the potential a given amount.* It is obvious that the greater the quantity needed to raise the potential by this amount, the greater will be the capacity for holding electricity. Symbolically, the capacity C is given by the equation

$$C = \frac{q}{\text{P.D.}}.$$

Here q is the quantity and $P.D.$ is the potential to which it is raised above some reference body.

In the electrostatic system the unit of electrostatic capacity of a conductor is the capacity which requires 1 electrostatic unit of quantity to change the potential by 1 electrostatic unit of potential. We must find the value of this unit from dimensional reasoning.

We found that the quantity

$$q = \sqrt{fr^2}$$

where f is the force between the two equal charges q and r is the distance between them. Thus dimensionally

$$q = \sqrt{ML^3T^{-2}}, \text{ in the electrostatic system.}$$

Again potential

$$\text{P.D.} = \frac{\text{Work}}{q} = \frac{ML^2T^{-2}}{\sqrt{ML^3T^{-2}}}, \text{ in the electrostatic system.}$$

Therefore

$$C = \frac{q}{\text{P.D.}} = \frac{q^2}{W} = \frac{ML^3T^{-2}}{ML^2T^{-2}} = L, \text{ in the electrostatic system.}$$

Therefore, the electrostatic unit of capacity has the dimensions of a length. Since q and P.D. in absolute electrostatic units are in

terms of the absolute C.G.S. system, the C.G.S. electrostatic unit of capacity is the cm.

That capacity has the dimensions of a length can be shown by considering the capacity of a sphere. The potential of a conducting sphere infinitely far away from any other charges was given as q/r (page 190), where r is the radius of the sphere. C therefore is

$$\frac{q}{\frac{q}{r}} = r, \text{ which is a length.}$$

In electromagnetic units, the unit of capacity is the capacity which requires one electromagnetic unit of quantity to change the potential by one electromagnetic unit of potential. It is seen that since the electromagnetic unit of quantity is a very large quantity and the unit of P.D. is a very small quantity, this unit of capacity must be a tremendous unit. In terms of the practical system, the unit of capacity is amperes times seconds divided by volts, or

$$\frac{\text{coulombs}}{\text{volts}} = \frac{Q}{V}.$$

This unit is called the farad. Therefore C (practical) equals

$$\frac{\text{coulombs}}{\text{volts}} = \text{farads}.$$

Now 10 coulombs equals an absolute E.M.U. and one volt $= 10^8$ absolute E.M.U. Therefore, one unit of capacity in absolute E.M.U. is equal to

$$\frac{10}{10^{-8}} = 10^9$$

practical units of capacity. Again the volt is $\frac{1}{300}$ of an absolute electrostatic unit of potential and the unit of quantity in the electrostatic system is

$$\frac{1}{3 \times 10^{10}}$$

absolute E.M.U. of quantity, therefore

$$C \text{ in E.M.U.} = \frac{\dfrac{3 \times 10^{10} q}{\text{P.D.}}}{3 \times 10^{10}} = 9 \times 10^{20}\, C \text{ in E.S.U.}$$

Thus, unit capacity in electromagnetic units is 9×10^{20} *cm.*

The practical unit of capacity, the farad, is so large a unit that it is impractical to use it. Most capacities used in experiments and industrial work are of the order of 1/1,000,000 of this. This leads to another unit. It is called the *microfarad;* that is, the microfarad is 10^{-6} farads $= 10^{-6} \times 10^{-9} = 10^{-15}$, absolute electromagnetic units. The microfarad is therefore 9×10^5 absolute electrostatic units. Or, the microfarad = 900,000 cm. In static electricity and for high frequency radio oscillations capacities are sometimes given in cms, otherwise capacities are given in microfarads, or micro microfarads, i.e., 10^{-6} microfarads or 10^{-12} farads or 0.9 cm.

81. CAPACITIES OF SPHERICAL, PARALLEL PLATE AND CONCENTRIC CYLINDRICAL CONDENSERS

Having defined the units and given their relative magnitudes, we may now turn to a study of capacities in various electrical systems. It will be seen at once from the definition we have given of capacity that capacity is a derived unit analogous in this respect to resistance in the electromagnetic system. Resistance was derived as the ratio between current and potential, the two fundamental quantities in current electricity. Resistance was shown to be a *characteristic of the particular system used.* Likewise, capacity is the ratio between quantity and potential which are the fundamental quantities of importance in the electrostatic system. As will be seen, this property is again a property of the conductors and their positions in space.

Let us calculate the capacities of various standard condensers.

(1) **Concentric Spheres.**—In the system shown in Fig. 68, there are two concentric spheres of radii A and B separated from each other and insulated, with all substances removed from between the plates, including air. The outer sphere of radius B is connected to the earth E. The inner sphere is connected to the battery or source of potential whose other terminal is also connected to the earth. In making the connection assume that we have placed a quantity $+ q$ on the inside sphere. A quantity $- q$ runs from the earth to the inside of the outer sphere. The potential of the sphere of radius A is by Chapter XIV equal to q/A; that throughout the space inside of B, due to the charge

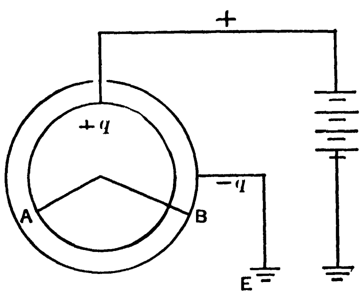


FIG. 68.—The Spherical Condenser.

on B, is $-q/B$. Thus the resultant potential difference between A and B is

$$\text{P.D.}_{AB} = \frac{q}{A} - \frac{q}{B}.$$

Since by definition

$$C = \frac{q}{\text{P.D.}}$$

Therefore

$$C = \frac{q}{\dfrac{q}{A} - \dfrac{q}{B}} = \frac{BA}{B - A}.$$

Thus it is seen that the capacity of a condenser made of concentric spheres is a function of its dimensions only. It is seen that the greater A and B, the greater is C, and the smaller $B-A$, the greater is C. It is thus seen that C depends on (1) the absolute size of the conductor, (2) the distance to the nearest body or bodies of opposite sign depending on the nature of the system. In the assumption made in the deduction of this equation all material was removed from between the spheres. Had different material substances been placed between the spheres the capacity would have been found to depend on the nature of the substance placed between the two spheres. Thus, it is seen that capacity finally depends on (3) the material between the two surfaces. This last item is an *experimental* observation originally due to Faraday. He found that the capacity of a spherical condenser changed with the material between the conductors. The constant which determines the value of this change in capacity is called the *dielectric constant* and is designated by the symbol D. For free space, it is unity. For air it has the value 1.00059. Thus for air D is nearly unity. For other substances it varies, being about 6.0 for glass, 4 for sulfur, and 2 for paraffin. Water has one of the highest dielectric constants among common substances. Its value is about 81. Liquid ammonia has a value which is higher than most liquids, being in the neighborhood of 16 at 14° C. It is to this high dielectric constant, which is also termed specific inductive capacity, that we owe the phenomena of electrolysis. This quantity D is identical with the D in the Coulomb law of force.

$$f = \frac{qq'}{Dr^2}.$$

We may define the term *dielectric constant* as the factor by which the force between two quantities q and q' in Coulomb's law of electro-

static attraction is reduced by the presence of a material surrounding the two bodies. It may also be defined as the factor by which the material between the plates of a condenser increases the capacity of the condenser.

The increase in capacity due to this dielectric constant depends on the fact that dielectric substances have in them small electrical dipoles. These dipoles are usually oriented in all directions. When a field is put on, the positive poles of the dipoles turn toward the negative terminal of the condenser and the negative poles turn toward the positive terminal of the condenser. Thus, as in magnetism, we have an orientation, in this case of *electrical molecular* dipoles. These take up the lines of force due to the field, and partially neutralize the charges collected on the surface of the condenser plates. They thus enable one to place more electricity per cm^2 on the surface before one increases the potential the same amount as in the absence of the dielectric. The reduction of the force is accomplished in the same fashion, since in the presence of a dielectric, the lines of force which would have gone from one body to the other, thus increasing the pull of one body on the other, are now in part neutralized by the material of the dielectric. The force is thus lessened.

Electrical dipoles or electrical doublets consist of a positive and an equal negative charge, or groups of positive and negative charges equal in amount, which are separated by a finite distance such that placed in a uniform electrical field they will have a torque exerted on them. The atoms or molecules are usually pictured as electrically neutral and having the center of positive and negative electrification of their constituent parts so completely overlapping that no external field is appreciable. Owing to the method of combination of two atoms to form a molecule, or owing perhaps to the fact that electrifications in the atoms are such that the negative charge is distributed over a finite number of electrons while the positive charge is concentrated at a point in the center of the atom, it often happens that there is not a complete superposition of the positive and negative fields. This means that the center of negative electrification is sometimes separated from the center of positive electrification by a small finite distance. The result is that the molecule or atom acts as a small electrical dipole which has a torque exerted on it in any electrical field. The action of an electrical field on a dipole is closely analogous to the action of a magnetic field on a bar magnet. The electrical dipole is characterized by an electrical moment, whose value is the product of the value of one of the equal charges and the distance between their centers. As stated, certain molecules in substances

like HCl, H_2O, $NaCl$, i.e., strong electrolytes, and also the weaker ones, have prominent electrical dipoles as a result of molecular structure. In many other substances the action of an external electrical field produces a displacement of the centers of positive and negative electrification in the atom due to the fact that the electron shell is distorted by the inducing field. Hence all atoms placed in an electrical field have an *induced* dipole moment so that all substances have a dielectric constant greater than unity. Some substances have these dipoles only in electrical fields, due to the inductive action of the field in producing dipoles, and others have dipoles due to the existence of permanently separated electrifications in the molecule. In the latter case we can typify this behavior by the molecule of HCl which consists of a hydrogen atom and a chlorine atom combined in such a way that the electron of the hydrogen atom in combination with the chlorine atom has been so displaced toward the chlorine that the chlorine atom acts as if it were negatively charged while the hydrogen atom appears to be positively charged. The dipole could be described by placing a positive charge separated from a negative charge by 0.217×10^{-8} cm, the charges being the charge on a chlorine ion and a hydrogen ion (i.e., minus and plus 4.77×10^{-10} E.S.U.), and the moment of this dipole is then given by 1.034×10^{-18} while the distance between the nucleus of the hydrogen atom and the nucleus of the chlorine atom varies between $(1.276 \text{ to } 1.34) \times 10^{-8}$ cm. It is seen therefore that in such a dipole the displacement of the centers of electrification is small and is not necessarily comparable with the distance between the two atoms. It is the action of these dipoles in an electrical field which accounts for the dielectric constants greater than unity.

It was by means of the spherical condenser that Faraday first measured dielectric constants of substances. If one include the dielectric constant of the material in the equation for the spherical condenser, one must write the equation for the capacity of such a condenser as

$$C = \frac{D(BA)}{B - A}.$$

(2) **The Parallel Plate Condenser.**—In the case of two parallel plates separated by a distance d, small compared to the linear dimensions of the plates, we can calculate the capacity very easily. In this case, Fig. 69, *the field is uniform;* that is, the lines of force run from one plate directly to the other normal to the surface of the plates except at the edges. (See page 192.) Since the number of lines of

force per cm^2 in this case gives the field intensity, in other words, the force on unit charge, we can assume if there are σ charges per cm^2 of surface that the field F equals $4\pi\sigma$. If there is a dielectric of value D between the plates,

$$F = \frac{4\pi\sigma}{D}.$$

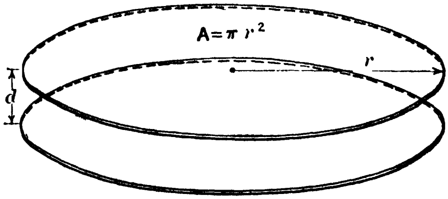


FIG. 69.—Parallel Plate Condenser.

Now in the last chapter, we showed that the potential difference P.D.$_{ab}$ is the work to move unit charge from b to a. Therefore,

$$-\text{ P.D.}_{ab} = \int_b^a \frac{(4\pi\sigma)}{D}dx = \frac{4\pi\sigma(a - b)}{D} = \frac{4\pi\sigma d}{D}, \text{ where } a - b = d.$$

If the area of the plates be A,

$$q = A \times \sigma.$$

Therefore

$$C = \frac{q}{\text{P.D.}} = \frac{DA\sigma}{4\pi\sigma d} = \frac{DA}{4\pi d}.$$

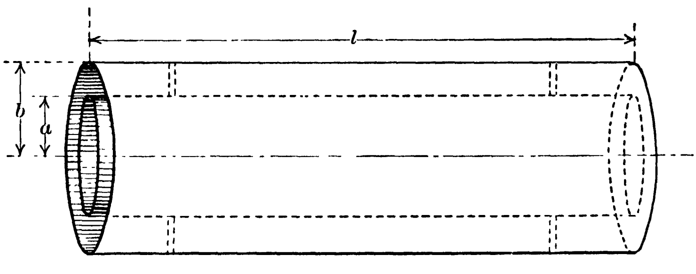


FIG. 70.—Cylindrical Condenser.

(3) The Cylindrical Condenser.—In this case we have two concentric cylinders of radii a and b, Fig. 70, the inner cylinder a being connected to one terminal of the battery, the outer cylinder b being grounded. Since the field is uniform except at its ends, for a cylinder in which the distance $b - a$ is small compared to the length l of the cylinder, the lines of force run radially from a to b and are symmetrically spaced about the axis of the cylinder. The field strength at any distance r from the axis of a cylinder carrying a charge q per cm

length is given by $2q/r$. (See Starling, "Electricity and Magnetism," page 127.) The potential difference

$$P_a - P_b = \int_b^a F dr = \int_b^a -\frac{2q}{r} dr = 2q \log \frac{b}{a}.$$

Since per unit length of the cylinder the charge on a is the charge of the system, and since b is earthed, and consequently $\text{P.D.}_b = 0$, therefore

$$\text{P.D.}_a = 2q \log \frac{b}{a}.$$

The quantity q of electricity per unit length of cylinder on a is equal to q. Therefore since $C = q/\text{P.D.}$ we have

$$C = \frac{q}{2q \log \frac{b}{a}}$$

as the capacity per unit length of a cylindrical condenser. Thus for a cylindrical condenser

$$C = \frac{1}{2 \log \frac{b}{a}}$$

per unit length. For a condenser of length l filled with a substance of dielectric constant D the capacity C is

$$C = \frac{Dl}{2 \log \frac{b}{a}}.$$

Condensers of this type are very easily made. For a small distance between a and b compared to the length, the capacity may be computed from this formula with a satisfactory degree of precision. Such condensers can thus furnish easily reproducible standards for all electrostatic capacity measurements. While the concentric spheres have less correction due to end effects than the concentric cylinders, the difficulty of turning out accurately constructed spherical condensers is too great to warrant their use.

82. USES OF CONDENSERS, ENERGY OF CHARGE, CONDENSERS IN SERIES AND PARALLEL

The ability of a condenser to hold electricity in large quantities at a low potential makes it very useful for many purposes. The uses of condensers are, among others, the following:

(1) An accumulator of electricity. (This is used in static machines and rectifiers for storing the charge until it is utilized.)

(2) As a potential multiplier. (The property of the condenser which leads to this use will be discussed later.)

(3) As a coupler of circuits. (For oscillating currents when it is unsatisfactory to make direct electrical connection with a circuit in which oscillation is taking place, use may be made of the induced oscillations set up in an insulated system through the inductive action of the condenser.)

(4) As a producer of oscillations. (It is frequently used in connection with an inductance for producing periodic variations of electrical intensity.)

(5) As a potential divider used in the measurement of high potentials with low range static instruments.

(6) For measuring small currents from the relation $i = C\frac{dV}{dt}$, in which C is the capacity and $\frac{dV}{dt}$ is the time rate of change of potential which can be measured.

Before taking up the various combinations of capacities, it may be worth while to calculate the energy of the charge on a condenser. By definition, dw, the work done in bringing up a quantity of electricity dq to a condenser against a potential P.D., is given by $dw = \text{P.D.} \times dq$, for by definition, $\text{P.D.} = \frac{dw.}{dq}$

Accordingly as we charge up a condenser by transferring to it a quantity dq successively we will find that the work required to add each succeeding dq is greater than the preceding one, owing to the fact that as each successive quantity dq is added the potential has been increased by a certain amount. Thus to determine the energy necessary to charge a condenser we must sum up the little increment of work $dw = \text{P.D.} \times dq$ for all values of q beginning at zero and going to the total charge on the condenser. This is done by the process of integration. Hence we can write

$$w = \int_0^q \text{P.D.}\, dq.$$

Since by definition

$$C = \frac{q}{\text{P.D.}}, \quad \text{P.D.} = \frac{q}{C},$$

and $w = E$ (the energy put into the charged condenser), we have

$$w = \int_0^q \frac{q\,dq}{C} = \frac{1}{2}\frac{q^2}{C} = \frac{1}{2}\text{P.D. } q = \frac{1}{2} C\,\text{P.D.}^2$$

Accordingly, when energy is stored on an ideal condenser the energy of charge is $\frac{1}{2}q$ P.D. or $\frac{1}{2}C$ P.D.^{2} This energy can be liberated in discharging the condenser either in the form of light, noise and heat, as in the spark, or as i^2r, heat in a conductor. Since the internal resistance of a charged condenser is very low, the discharge of a large capacity at a high voltage will give instantaneous currents of high value. In this way, Dr. Anderson at Mount Wilson Solar Observatory " explodes " fine wires of substances and gets instantaneous temperatures of 20,000° C, i.e., about four times as high as that of the sun.

If a condenser of capacity C_1 at a potential V_1 be connected to a capacity C_2 at a potential V_2 by wires, the potentials will equalize and electricity will be transferred from one to the other. It is of interest to calculate the energy after transfer.

Let V and C represent the potential and capacity of the combined system. If V_1 is greater than V_2 the quantity transferred q will be

$$q = (V_1 - V)C_1 = (V - V_2)C_2,$$

and

$$V = \frac{V_1C_1 + V_2C_2}{C_1 + C_2}.$$

Before connecting the condenser, the energy was

$$E_{12} = \tfrac{1}{2}C_1V_1^2 + \tfrac{1}{2}C_2V_2^2.$$

After connecting the condenser the energy was

$$E = \tfrac{1}{2}(C_1 + C_2)\left(\frac{C_1V_1 + C_2V_2}{C_1 + C_2}\right)^2 = \frac{1}{2}\frac{(C_1V_1 + C_2V_2)^2}{C_1 + C_2}$$

$$E_{12} - E = \tfrac{1}{2}C_1C_2\frac{(V_1 - V_2)^2}{C_1 + C_2}.$$

As this energy is positive, E_{12} is greater than E. The amount of energy computed above is thus lost when q goes from V_1 to V_2. This loss must go to heating the wires during the flow of current, just as when water flows from one tank to another because of a difference in gravi-

tational potential its energy is consumed in friction and heating the pipes. The heating of the wire on the sharing of charge between two condensers, or the discharge of one condenser, may take place in one of two ways. If the resistance is high the current will flow slowly and the whole energy of the current will go to raising the temperature of the conductor. If the resistance is small, exactly as would be the case if two large tanks of water were connected by a large pipe which was suddenly opened, the charge would oscillate backward and forward, each oscillation heating the wire by a small amount. The net result would be the same, and the two ways in which the heating may take place are adequately described in Chapter XXV.

In the use of most condensers there is another source of energy loss which is not included in the treatment above which applies to ideal condensers only. This loss is due to what is termed hysteresis in dielectrics. Beyond a slight conduction through the dielectric which can be neglected as far as energy loss is concerned, putting a field across a condenser containing, say, paraffined paper, orients the minute molecular dipoles in the dielectric. This orientation of the dipoles, just as in the case of magnetism, must be done against the resisting cohesive forces. Thus work is done, and this goes to molecular agitation or heating. Hence in an alternating field where the molecular dipoles are twisted first in one sense and then in the other, the dielectric is heated and energy is lost. Such losses are not negligible and much work has been done to reduce the losses by finding proper insulators. This orientation will often manifest itself in direct currents by the presence of charges for some time after the condenser was discharged the first time. These are due to the release of charges bound on the surface of the metal by the oriented dipoles, as the latter return to their normal positions, due to heat motions. Such charges may however also be due to mobile ions in the dielectric which act in a similar way, as recently shown by Joffé, and both effects are present. It will be seen that the molecular *magnetic* dipoles, or magnets, produce precisely the same phenomena due to the resistance to orientation in a magnetic field (see Chapter XVIII), and we speak of magnetic hysteresis.

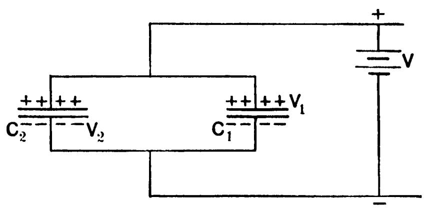


FIG. 71.—Condensers in Parallel.

It is of importance to calculate the values of the capacities for various combinations of capacity. This will be done in what follows.

Capacities in Parallel.—In Fig. 71, two condensers C_1 and C_2 are connected in parallel. The potential V_1 across C_1 and that of V_2 across C_2 are both equal to the potential V across the line. By definition

$$V_1 = q_1/C_1,$$
$$V_2 = q_2/C_2,$$

and

$$q = q_1 + q_2.$$

Therefore since $q = CV$ (where C is the total capacity)

$$q = CV = C_1 V_1 + C_2 V_2;$$

and hence

$$C = C_1 + C_2.$$

Capacities in Series.—If we have two capacities in series across a battery, as shown in Fig. 72, the capacities having the values C_1 and C_2, we see that from the laws of electrostatics the quantity of electricity q_1 on C_1 must equal the quantity of electricity q_2 on C_2. These each are equal to the quantity q of electricity on the combined capacity C, of C_1 and C_2. That this must be so is seen from the following considerations.

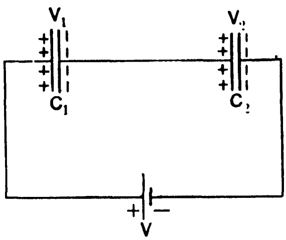


FIG. 72.—Condensers in Series.

If a plus charge be placed on C_1, negative electricity will be drawn from the insulated section containing one plate of C_1 and one plate of C_2. The amount bound by the positive charge on the left-hand side of C_1 must be equal to the positive charge on C_1. On the other hand, since initially the insulated section was neutral, the equal quantity of positive electricity will be left on the left-hand side of C_2. The latter will in turn bind an equal quantity of negative electricity on the right-hand plate of C_2. Thus the charges on the condenser C_1, on the condenser C_2, and on the condenser made of the left plate of C_1 and the right plate of C_2 must all be equal. Furthermore, since there is no current flowing, the potential V of the battery must be distributed in two parts—one across the condenser C_1, and the other across the condenser C_2. In equation form this is expressed by

$$V = V_1 + V_2.$$

Therefore

$$\frac{q}{C} = \frac{q_1}{C_1} + \frac{q_2}{C_2}$$

and

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}.$$

Thus

$$C = \frac{C_1 C_2}{C_1 + C_2}.$$

Furthermore

$$q_1 = C_1 V_1 = q_2 = C_2 V_2.$$

Therefore

$$\frac{C_1}{C_2} = \frac{V_2}{V_1}.$$

The last equation says that the fall of potential across the two condensers C_1 and C_2 is in the inverse ratio of their capacities. We can accordingly utilize this principle in the measurement of potentials with electrostatic voltmeters. If we have a total fall of potential across two condensers C_1 and C_2, by placing our voltmeter across C_2 and making its capacity large compared to C_1, then the reading of the potential across C_2 will be a simple fraction of the potential across C_1 and C_2. This follows from the fact that

$$\frac{C_2}{C_1} = \frac{V_1}{V_2},$$

whence

$$\frac{C_2 + C_1}{C_1} = \frac{V_1 + V_2}{V_2} = \frac{V}{V_2}.$$

If now we wish the potential across C_2 which is V_2 to be $1/x$ times as great as the total potential V, all that need be done is that the capacities C_1 and C_2 be adjusted in the following fashion. From the above equation

$$\frac{V}{V_2} = \frac{C_1 + C_2}{C_1} = x,$$

and thus

$$C_2 = C_1(x - 1).$$

This means that, if the capacity C_2 be made $x - 1$ times as great as C_1, the fall of potential across C_2 multiplied by x will give the potential across the circuit. This device is frequently used for distributing potential drops and for measuring high potentials with instruments of a low range.

Another use of capacity is in multiplying the effects of small quantities of electricity. Assume we have a gold leaf electroscope, G,

Fig. 73, and let us fasten on the insulated electrode a large disc A, which forms the base of a parallel plate condenser. Separated from this plate by thin, accurately equal, insulators, we must place a second plate B of the parallel plate condenser which is earthed and which is provided with means for removing it easily. If now a charge be placed on the lower plate A, the upper plate B, and A, again form a condenser. After charging the system, assume that the contact to the outside be broken, and that the upper plate be removed to a considerable distance. Then the charge which had been placed on A and had charged the capacity of AB, which we will designate as C_1, to a potential V_1 will reside on a system consisting of the plate A alone, of capacity C_2 which is much smaller. If the ratio between C_1 and C_2 is known, then the ratio between the quantities $q = C_1V_1 = C_2V_2$ will also be known. Thus:

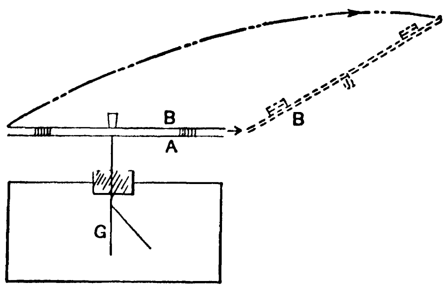


FIG. 73.—Electrostatic Potential Multiplier.

$$\frac{V_2}{V_1} = \frac{C_1}{C_2},$$

which says that the deflection given by the electrostatic voltmeter when the upper plate has been removed will be to the deflection when the upper plate is present as the capacity when the upper plate is present is to the capacity with the upper plate removed. In this way, we can make the potential of a dry battery, which is very much too small to cause a gold leaf electroscope to deflect, serve to give a measurable deflection on a gold leaf electroscope. By accurately determining the capacities, this multiplication can be made as accurate as is desired.

CHAPTER XVII

ELECTROMAGNETICS

83. DEFINITION OF CURRENT AND WORK DONE IN MOVING A CONDUCTOR IN A FIELD

In Chapter V, we started with two definitions of current in the electromagnetic system. The first definition of current was used in determining the strength of the current and in subsequent applications.

We now come to a very useful application of the second definition of current. *Unit current is that current which flowing in unit length of a conductor perpendicular to a unit magnetic field causes the conductor to experience a force of* 1 *dyne.* Symbolically, this is expressed in terms of the equation

$$f = i_a l H.$$

If i_a = one absolute electromagnetic unit of current,

$l = 1$ cm.

$H = 1$ gauss,

then

$$f = 1 \text{ dyne}.$$

This equation $f = i_a l H$ is the basis of calculating forces on all conductors carrying currents in magnetic fields. In fact, it was the basis of the calculation underlying the theory of the moving coil galvanometer in Chapter IX.

If we have a conductor of length l lying perpendicular to a magnetic field of strength H and pass a current i_a through it, a force will act on the conductor and cause it to move. Since f (the force) is given by $f = i_a l H$, then the work $(W) = i_a l H x$ when the conductor moves x cms in the field.

$lx = A$ where A is the area swept out by the conductor, hence $W = i_a A H$. AH is the total number of lines of force cut, or *the total flux*. If we designate *flux* by the letter ϕ, then $W = i_a \phi$. This says that the work done, when a conductor carrying a current cuts a mag-

netic field, is the current multiplied by the total number of lines of force cut. This leads to certain useful applications.

84. FIELD IN AN INFINITELY LONG SOLENOID

Owing to the difficulty of measuring magnetic moments and field strengths it would be very valuable to have a means of obtaining uniform reproducible fields of known strength whose measurement or calculations consists in a substitution of the value of an easily measured current in an equation. Such an aid actually can be found through the study and use of the field in an infinitely long solenoid, which for practical purposes can be taken as a coil whose diameter is small compared to its length.

Consider a coil of n_1 turns of wire carrying a current i_a. Then take a magnetic pole of strength m and move it once around the wires of the coil. Each wire in this coil has cut the $4\pi m$ lines going out of the magnetic pole once. Thus when we carry a pole of strength m around a coil of n_1 turns, carrying a current i_a, each wire is cut by $4\pi m$ lines of force, and from the above, the work done, $W = 4\pi i_a m n_1$.

If we have a long coil, by this we mean a solenoid whose diameter is small compared to its length, we have a very interesting case. If one studies the direction of the lines of force encircling each wire for two adjoining turns. one would observe that these lines of force annihilate each other except parallel to the surface of the coil on its inside and on its outside. The field, therefore, represented in such a section would consist of lines of force parallel to the axis of the coil running in one sense inside the coil and in the other sense outside the coil. There is thus no loss of lines of force through the sides of the coil, and all lines of force originating in the coil run through the length of the coil emerging at its ends. As a result of this, the lines of force distribute themselves uniformly inside the coil and run parallel to its axis. One has, therefore, a uniform magnetic field down the center of the coil. If now the coil with n_1 turns were bent into a circle of radius r so that the two ends were together, we would have the cir-

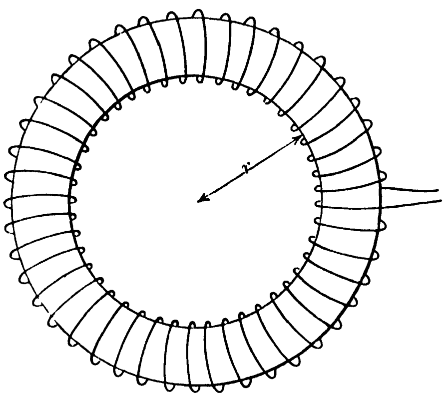


FIG. 74.—Coil Bent into a Circle.

cuit of the lines of force continuous inside the coil, as none would emerge, see Fig. 74. Thus the lines of force give a uniform field of length $2\pi r$ which is the length of the coil. The work done in moving a pole of strength m once around the circuit is then $W = 2\pi r H m$. If on the other hand the pole m had been carried around each wire once separately, the work would have been $W = 4\pi n_1 i_a m$. The work in both cases would have been the same, as in each case m was carried around each wire once. We can therefore write

$$W = 4\pi n_1 i_a m = 2\pi r H m$$

or

$$H = \frac{4\pi n_1 i_a}{2\pi r}.$$

Now assume that the radius r of the circle into which the large coil

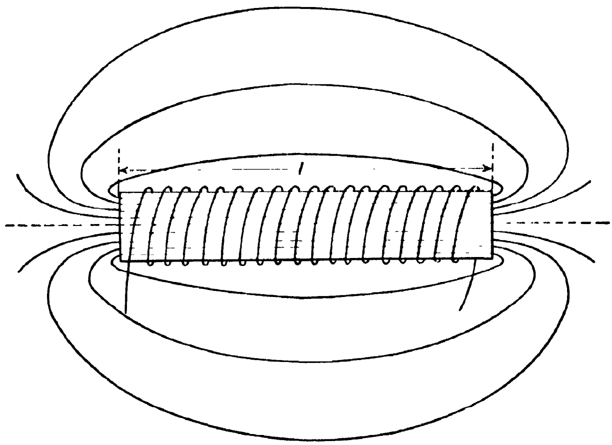


FIG. 75.—Magnetic Field of a Solenoid.

was bent is *very large*. A limited section of the coil is nearly a straight line, and the field except near the ends is H, which is the same as that above. Therefore we can set the field in such a section as given by $H = \dfrac{4\pi n_1 i_a}{2\pi r}$. In this section, however, we have not n_1 but n turns and a length for these n turns of l cm. Since for a uniform coil $\dfrac{n_1}{2\pi r} = \dfrac{n}{l}$, we may write

$$H = \frac{4\pi n i_a}{l},$$

where l is the length of the section. Thus the field in a straight section of a solenoid the length of which is great compared to its diameter is approximately $H = \frac{4\pi n i_a}{l}$ and if i is in amperes we can write $H = \frac{4\pi n i}{10 l}$ for the field in a long solenoid.

This is, however, only an approximation which is the more correct the more nearly the ideal conditions above are approached. For ratios of length to diameter the order of 10 to 1 the equation is good to a fraction of a per cent in the center of the coil. The deviations are greatest at the ends of the coil, as shown in Fig. 75, where the lines of force begin to diverge, and at distances from the wires of the order of the diameters of or distances between the wires. Corrections for these effects can be found in the " Bulletin " of the Bureau of Standards and certain advanced texts. Coils of this design are of great value, as they enable us to produce uniform and calculable magnetic fields, which are especially important in the study of electromagnetism.

85. INTERACTIONS BETWEEN FIELDS AND CONDUCTORS CARRYING CURRENTS

The study of the interaction between current and magnetic field is essential in order to be able to predict the direction of motion resulting from such an interaction. The study of these laws is often facilitated by the use of certain rules known as the dynamo and motor rules. These are hard to remember, easily confused, and a simple device for enabling one to determine the direction of motion is much to be preferred. The method to be outlined in what follows is one which is very simple. It requires memory only of the *right-hand rule*, which gives the direction of the magnetic field circling a wire carrying a current.

Consider the wire represented by i in Fig. 76*a*. This is in a uniform magnetic field H at right angles to the wire. The field about the wire is indicated by the arrows. In cross section we would see the fields as indicated in Fig. 76*b*. It is seen there that on the lower side of the wire, the lines of the field about the wire are in the same direction as the field of the magnet. On the upper side, the lines of force are contrary in sense. These result in a field of the form illustrated in Fig. 76*c*. This occurs since two superposed fields in opposite senses when added give a weaker resultant field. The field H above is therefore weakened and is represented by less lines per unit area.

It is seen that the lines of force on the lower side of the wire are crowded together. On the upper side, they are somewhat apart. Lines of force as imagined by Faraday act like stretched elastic rubber bands which exert stresses as a result of distortion. These stresses have components perpendicular to the general trend of the lines. Since, on the lower side of the figure, the number of lines compressed together is greater than on the upper side, the resultant force on the wire would be to move the wire upward relative to the magnet. In general on plotting the lines of force about the conductor where the lines of force go in the same direction as the field, the force is greater than on the side on which they go in the opposite direction. Motion will always be from the side of greater number of lines to the lesser.

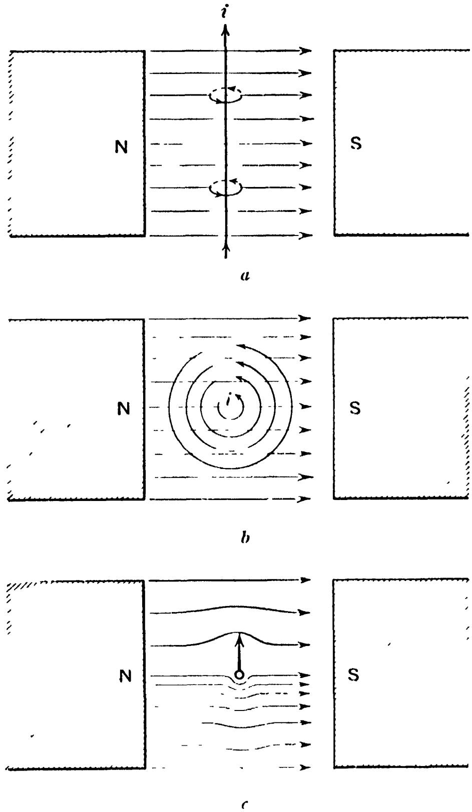


FIG. 76.—Determination of Direction of Motion of a Conductor Carrying Current in a Magnetic Field.

We may now apply this rule to several simple cases.

(*a*) Consider two wires with the currents flowing in the same direction. In what way will these wires react in reference to each other? Fig. 77*a* illustrates the direction of the fields about the two wires. It is seen that outside of the two wires the forces produced by the one wire and the other are in the same direction whereas between the wires the forces annihilate each other. If the currents in the two wires were equal, there would be no field in the neighborhood of the center as one approaches one wire from the other. Since the field increases as we approach the wire, there will be a few lines of force and those will increase in number as the wire is approached. The result will be that the wires will act as if they were being drawn together, that is, they will attract each other. Again, if the current

flow upward in one wire and downward in the other, as shown in Fig. 77*b*, the lines of force will add up in the space between the wires while in the region surrounding both wires they will annihilate each other. Thus there will be a crowding of the lines of force between the wires and the wires will appear to repel each other.

The forces can be easily calculated as follows. If we assume that

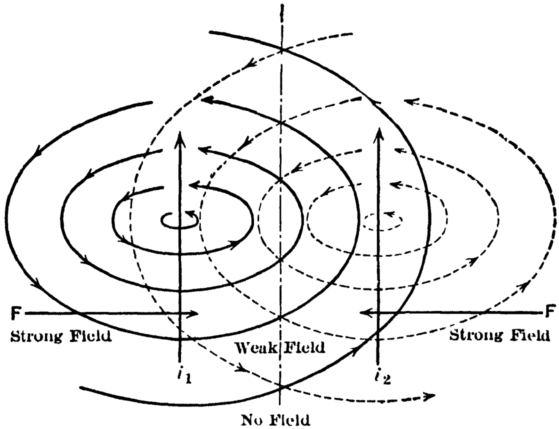


FIG. 77*a*.—Fields about Parallel Conductors; Currents in the Same Direction.

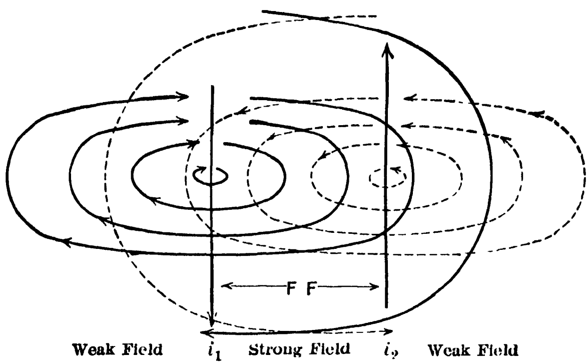


FIG. 77*b*.—Field about Parallel Conductors; Currents in Opposite Directions.

the wires are infinitely long, we may apply Biot and Savart's law. Such an assumption is realized to a fairly good degree of approximation in cases where the distance between the wires is about one-tenth the length of the wires. At r cm from a wire a, with current i_a, its field is $H = \frac{2i_a}{r}$ from Biot and Savart's law. The force of H on wire b with a current i_b is Hi_bl, where l is the length. Therefore, the force

between wires a and b is one of attraction or repulsion which is given by

$$f = \frac{2i_a i_b l}{r}.$$

This force is in dynes if i_a and i_b are in absolute E.M.U.

Another interesting combination is what is known as the Faraday disc. It is also called Barlow's wheel. This consists of a copper disc * mounted on an axle, shown in Fig. 78. This is placed between the poles of a powerful magnet N–S. The lower end of the disc dips into a trough of mercury Hg while the axis and the trough of mercury are connected to the terminals of a source of potential. The current then flows into the axle down through the mercury and back to the battery. This current is acted on by the magnetic field causing the disc to rotate in a sense determined by the direction of the field and the direction of the current. The force then acts tangentially on the wheel causing it to turn. This constitutes one of the first electric motors; it is, however, inefficient because of the effect of eddy currents which we will discuss in a later chapter. The inefficiency of this motor led Faraday to conclude that the practical application of his discoveries would never be of any value. Not very long after these statements, a more practical form of motor was devised.

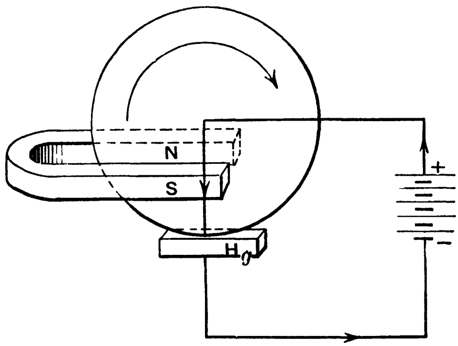


FIG. 78.—Barlow's Wheel.

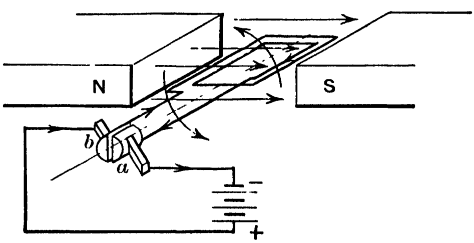


FIG. 79.—Simple Motor.

This simple motor which we may now discuss is shown in Fig. 79.

A coil of wire is placed between the poles of a magnet NS. One end of the coil of wire is connected to the half of the split ring a,

* The Barlow wheel will operate in the absence of the magnetic field introduced from outside. This seems paradoxical at first, until it is recognized that the field produced by the circuit $ABCD$ carrying the current to the wheel itself can produce a field which will act in the same way as the field of the magnet.

the other end is connected to the half of the split ring *b*. The segments of this split ring are insulated from each other. The terminals of the battery are connected to the two terminals which are at the two ends of the diameter of the ring. When the current flows through the coil, the wires which are normal to the lines of force of the field will have forces exerted on them urging the one wire up and the other wire down, provided the current and field are properly arranged. The forces on the two wires produce rotation in the same sense and there is a couple acting to cause the motor to turn. The wires turn in such a fashion that as they reach a position of equilibrium (when the plane of the coil is perpendicular to the field), in which the field acting on the wires reverses in direction with regard to the flow of current in the wire, the current is also reversed in direction by the contacts striking the next segment of the split ring. Thus the two wires are caused to rotate continuously about their axis. It is seen that when the coil is perpendicular to the field the force is zero. Accordingly, unless the inertia of the rotating system carries the coil past this neutral position the coil will cease to rotate. The motor coil is then said to be on a dead center and it must be displaced from this before the field is turned on to start the motor. This action is very well illustrated in the cheap commercial toy motors in which friction is high and the inertia of the armature is so small that the coil readily finds itself in a position of dead center. The energy consumed and the considerations of the torque produced by the motor will furnish material for another chapter.

Still another example of the force between wires is illustrated in the case of a loose coil of spring wire suspended vertically which carries a current and dips into a cup of mercury. Two succeeding turns of the wire carry the current in the same sense. This results in the field produced in the two turns being such that the turns attract each other. If the spring be suspended so that one end dips into a trough of mercury and the trough of mercury be connected to a battery whose other end goes to the further end of the coil, then as soon as contact is made a current flows through the coil. The coil contracts against its elastic forces, breaking the contact. On breaking contact the force is removed and the coil again makes contact. The result will be that the coil oscillates back and forth as long as the current flows through it. This oscillating coil is depicted in Fig. P-26, page 387.

There are beside this a large number of other examples, and the mechanics of each of these examples can be worked out very satisfactorily by means of the rules laid down.

CHAPTER XVIII

MAGNETIC PROPERTIES OF MATERIALS

86. THE MANNER OF REGARDING A PIECE OF IRON IN A MAGNETIC FIELD

As was stated in the second chapter substances which show magnetic properties are either permanently or temporarily magnetized in a magnetic field. They show various degrees of retentivity. Those with a high degree of retentivity are the permanent magnets; those with a low degree are temporary magnets. The magnetization of a specimen depends on the magnetic field and at first increases slowly, then rapidly as the field increases, finally increasing more and more slowly, eventually reaching saturation. The explanation of the latter effect was made on the basis of the assumption of elementary magnets.

In the last chapter we saw that in a solenoid one can get a uniform magnetic field with lines of force parallel to the axis of the coil. Let us consider more closely what happens when iron or steel is placed in such a field. The apparent result is an increase in the number of lines of force over what there was when there was no field acting. This increase in the number of lines of force may be treated mathematically in *two ways.* The two modes of looking at the phemomenon are both useful for different applications to magnetic problems. The first point of view regards the increase in the lines of force from the point of view of the *creation of new magnetic poles in the iron.* The second point of view discusses the increase of lines of force as an *increase in the number of lines of force threading the iron.* It really considers the iron as a magnetic conductor.

Let us take the first point of view. Assume that we have a piece of iron of length l. It acts like a magnet of pole strength m and moment M. The magnet of pole strength m creates new lines of force. This magnet is, however, produced by the field H already existing in the coil. If the magnet exists only as long as the field exists (which is the case for soft iron) the magnetization is not permanent and may be proportional to the field. This is the simplest case to consider.

The new poles contribute $4\pi m$ lines of force. The density of flux where iron or magnetic materials are concerned is called the *induction* and is represented by the letter B. Thus we can write for a piece of soft iron in a field H:

$$B = H + \frac{4\pi m}{A}.$$

Here A is the area of cross section of the iron normal to the field, and the $4\pi m$ represents the lines due to the new pole. If we multiply the top and the bottom of the expression $\frac{4\pi m}{A}$ by l, the length of the magnet, then

$$\frac{4\pi ml}{Al} = \frac{4\pi M}{V},$$

where Al is the volume V of the magnet and ml is the magnetic moment M. Thus

$$B = H + \frac{4\pi M}{V} = H + 4\pi I.$$

I is a new quantity defined as the intensity of magnetization. It is the ratio of the magnetic moment to the volume of the magnet. Now I is produced by the field. Thus

$$I = \frac{M}{V} = \kappa H.$$

Here κ is a constant, over a small range of values of H, called the susceptibility. It measures the susceptibility to magnetization by the field H. Therefore

$$B = H + 4\pi\kappa H$$

and

$$B = H(1 + 4\pi\kappa).$$

From the other point of view, we regard the iron as a magnetic conductor and simply state that more lines of force flow per unit area taken perpendicular to the iron than would flow through the same space were the iron absent. These so-called magnetic conductors have a greater *permeability* than the air. This is stated in equation form by writing that $B = \mu H$, where μ is a number greater than 1 which gives the amount by which H inside the conductor must be multiplied to give the induction B. For empty space μ equals 1, while for iron it attains values of 1000 or more. For air μ is very nearly equal to 1, having the value of 1.0000004. We can therefore

tivity because of resistance to the orientation or rearrangement of the molecular magnets. In substances where there is no retentivity these rearrange themselves readily for there is no friction to keep them from doing so. With a high retentivity the arrangement of the molecules depends entirely on the work done against the friction. Thus the hysteresis loop above represents the work lost as heat in magnetizing and demagnetizing the iron. The heat comes from the frictional action of the molecular magnets. Fig. 82 shows the hysteresis loops for soft iron and steel. As is seen, the area under the steel loop is much greater than under that of soft iron.

88. THE MAGNETIC BEHAVIOR OF ATOMS, PARA-, DIA-, AND FERROMAGNETISM

We now come to a more careful scrutiny of magnetism. If we place small pieces of different types of matter in a *strongly divergent* magnetic field (that is, near a conically shaped pole piece of a magnet) it will be observed that certain of these substances are attracted to the magnetic pole while others are repelled. By a divergent field in contradistinction to a uniform field is meant a field where the lines of force rapidly separate and do not run parallel. In such a field the force changes rapidly with the distance from the pole. Those attracted are elements like iron, nickel, cobalt, Al, Cr, and Mn, while those repelled are substances like copper, bismuth, silver, etc.

Again, if we place oblong pieces of these substances in a *uniform field* the iron-like substances will set themselves parallel to the field so that as many lines of force as possible run through the body. The others will set themselves at right angles to the field trying to avoid the passage of magnetic lines. The bodies which are *attracted in the divergent field and set themselves parallel to the uniform field are called paramagnetic bodies.* Among the paramagnetic bodies, we distinguish some which show a feeble activity, e.g., Al, Cr, Mn, as against the powerful activity of Fe, Ni and Co. The latter bodies are termed ferromagnetic bodies. It seems that ferromagnetism is a particular manifestation of paramagnetism as yet little understood. Above certain temperatures ferromagnetic bodies lose this extreme property and become paramagnetic.

Those bodies which are *repelled by the divergent field and set themselves perpendicular to the uniform field are called diamagnetic bodies.*

The explanation of the two actions above is very easily seen if we assume in the case of the orientation of the bodies parallel or perpendicular to the field that the diamagnetic body tries to have *less* lines of force passing through it than pass through air, while the para-

magnetic bodies tend to *concentrate* the lines of force. The same is true if we look at the diamagnetic body in a divergent field. It is repelled because less lines of force are able to thread it than thread the space in its absence. The paramagnetic body acts as it does because *more lines* of force try to pass through it than pass through empty space. Considering forces as stretched elastic rubber bands, make these actions obvious in Figs. 83 and 84 which represent dia- and paramagnetic bodies respectively in a divergent field, and in Figs. 85 and 86 which represent oblong pieces of dia- and paramagnetic bodies

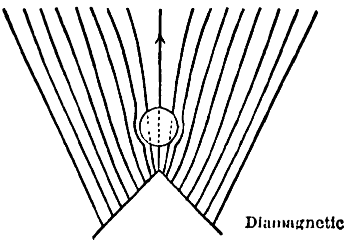


FIG. 83.—A Diamagnetic Body in a Divergent Field.

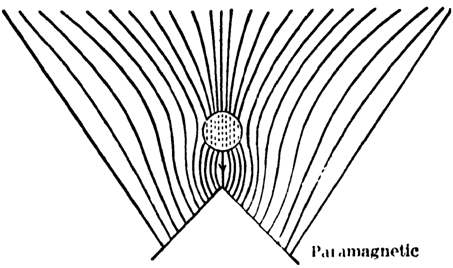


FIG. 84.—A Paramagnetic Body in a Divergent Field.

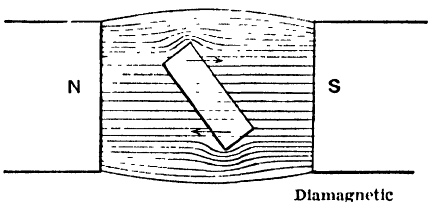


FIG. 85.—An Oblong Diamagnetic Body in a Uniform Field.

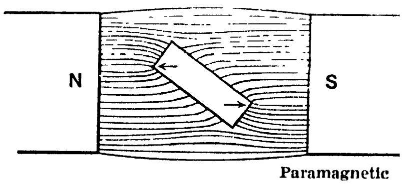


FIG. 86.—An Oblong Paramagnetic Body in a Uniform Field.

respectively in a uniform field. All this may be summed up in the sentence which says that for paramagnetic bodies μ is greater than one, while for diamagnetic bodies μ is less than one.

We have explained the fact that μ is greater than one by assuming that paramagnetic bodies have small magnets which align themselves in the magnetic field and this alignment changes the otherwise inert body into a magnet, temporarily or permanently. Thus, the increased permeability is explained.

In the case of the diamagnetism, one may ask how the permeability can be less than unity. The only explanation we have is that a magnetic field is induced in the body when the external field H is put on in such a direction as to oppose the exciting field H. That such an

action must follow comes directly from our modern understanding of atomic structure. Atoms consist of a central positive sun, *N*, Fig. 87, the nucleus, with negative electrons, such as *E*, describing orbits about it. An electron moving about a nucleus constitutes a small gyroscope. When a magnetic field *H* is placed in the neighborhood of the atomic gyroscope, the electron in its orbit which constitutes a current suffers a torque or force *F* about the axis *ANB*. The torque acting on a gyroscope causes the gyroscope to execute what is known as a precessional motion. This precessional motion causes the plane of the orbit to describe an orbit (indicated by the motion of the point *O* out of the plane of the paper towards *P*), in such a fashion that we have the electron moving partly so as to create a current producing a field parallel to the magnetic field *H* which caused the disturbance but opposite in sense to *H*. There is, as we shall see in a later chapter, a law which says that when a circuit is so disturbed by a magnetic field as to have a current set up in it the current set up will flow in such a sense as to counteract the magnetic field producing it. If this were not so we could get perpetual motion. Our little molecular gyrostat in the magnetic field is thus caused to precess so that the electrical gyrostats of the atoms set up a field opposing the inducing field. This causes the diamagnetic behavior. All substances should show diamagnetism and they *apparently* do. In the ferromagnetic and paramagnetic substances, however, the ferro- and paramagnetic effects so far outweigh the diamagnetic effects that the latter are not noticeable.

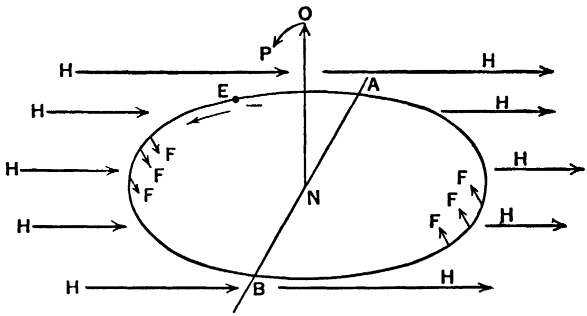


FIG. 87.—Diagram Illustrating the Nature of Diamagnetism.

89. THE EXISTENCE OF AMPÈRE'S MOLECULAR MAGNETS, THE MAGNETON AND THE QUANTIZATION OF ATOMIC MAGNETS

When Ampère, who was one of the early protagonists of the molecular and atomic theory, studied the magnetic fields due to currents he suggested that the molecular magnets in the magnetic atoms or molecules were due to small circular electrical currents in the atoms. With the discovery of the electron and the Rutherford picture of the

atom as a group of electrons moving about a central sun or nucleus in orbits some hundred years later the plausibility of the real existence of these Ampèrian molecular magnets was established. Two investigators attacked the problem from different angles. S. J. Barnett stated that if one accelerate a piece of metal some of the atomic gyrostats causing diamagnetism must be set into precessional motions due to the very small mechanical torque placed on such atoms as were properly oriented. Calculation showed that such a precession should cause the metal to become magnetized. Ferromagnetic substances were spun in a space free from magnetic fields so that their permeability would increase the minute effects to be expected from the very small magnetization coming from accelerations that were possible experimentally at the highest mechanical speeds of rotation, for our mechanical speeds and accelerations are minute compared to such quantities in atoms. The experiments consuming many years were finally successful but gave a value about half that to be expected by theory. At the same time Einstein and de Haas suggested that by properly applying a magnetic field to a diamagnetic body the precessional motion should set up a minute mechanical force which should manifest itself. By applying a periodic external field which synchronized with the period of oscillation of a suspended piece of metal the mechanical system experienced a feeble periodic torque. Thus the minute mechanical effect could be many times multiplied and easily detected by the motion of the metal. They actually demonstrated the existence of the torque. Here again the numerical values were less than the theoretical values by something near a half. The explanation of this is outstanding, but it seems that a solution of the difficulty is near.

By studying the paramagnetism of substances in solutions Pierre Weiss arrived at the conclusion that paramagnetic bodies were composed of combinations or multiples of a magnet of unit strength or moment. This elementary magnet representing what Weiss believed to be the elementary molecular magnetic moment he called the *magneton*, and he assumed it to be the expression of the unit Ampèrian current causing paramagnetism. When the structure of atoms became clearer, Bohr, whose pioneer work had done much to develop our knowledge, concluded that there was a different elementary unit of magnetism. He deduced the value of this unit theoretically from his theory of atomic structure. The moments in each case depend on the number of electrons and the types of orbits they are in. He also concluded that in the vapor of silver there must be one such magneton per atom and that these atmoic unit magnetic dipoles must in a field

orient themselves (contrary to the existing belief which assumed random orientation of the axes of the dipoles) with their axes in only two directions. Either the moments would be in the direction of the field, or they would be oriented at 180° with the field, i.e., they would set their north poles towards the north pole of the field.

Between 1921 and 1924 Stern and Gerlach passed a beam of silver atoms perpendicular to a very inhomogeneous magnetic field in a high vacuum. In such a beam the silver atoms traverse the field in some 10^{-4} seconds and suffer no disorienting collisions. If there is no field

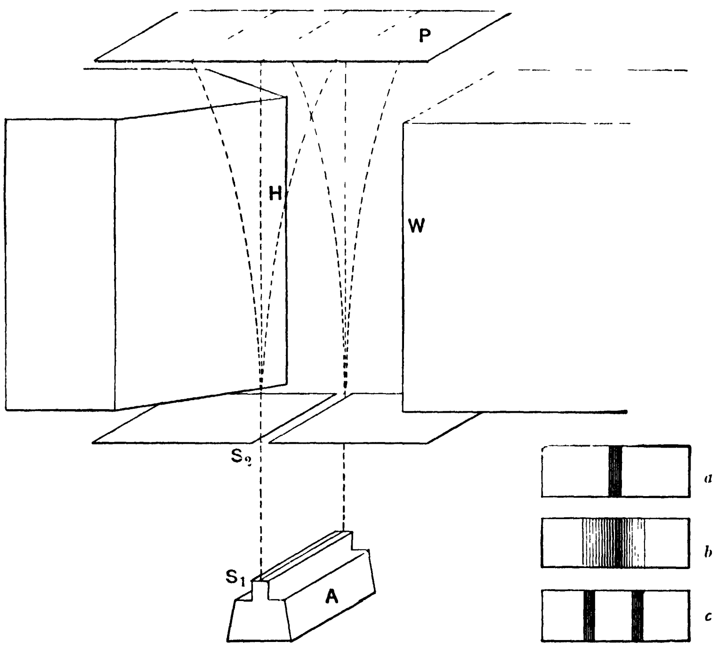


FIG. 88.—The Stern-Gerlach Experiment.

with the apparatus sketched in Fig. 88, where A is the oven with evaporating silver, S_1 and S_2 are slits to give a uniform beam, HW is the field strong along the knife edge H, weak along the pole piece W, the deposit on the glass plate P should consist of a sharply defined image of the slits as the atoms move in straight lines, Fig. 88*a*. If the molecular magnets are oriented in all directions as usually assumed, one might expect the field to produce a diffuse band wider than the slit image whose density should vary in some manner indicative of the orientation of the dipoles do they exist, Fig. 88*b*. If Bohr's picture is correct the single image of the slit in the absence of the field should

in the field be broken up into two sharply defined images to the right and left of the undisplaced image, as in Fig. 88*c*. This would occur because the dipoles with axes parallel to the field and N poles towards the south pole of the field are attracted in a divergent field while those with reversed moments are repelled. The separation if the gradient of the divergent field is known will then give the value of the atomic moment and decide between the Bohr and the Weiss magnetons. The experiments showed the type of behavior predicted by Bohr and the two separate images were observed. The value of the Bohr magneton was found from these measurements as predicted, and later theoretical work established the true relation between the Bohr magneton and the Weiss magneton. Other atoms were studied and some showed moments while others did not. In most cases the results fulfilled the predictions of the Bohr theory. Thus we see that essentially the magnetic behavior of substances is explained by Ampère's molecular currents. In the light of the modern view of atomic structure the existence of residual permanent moments in the atoms can in a measure be predicted as a function of their structure. The peculiar phenomenon known as ferromagnetism has however not been satisfactorily accounted for to date.

CHAPTER XIX

THE MAGNETIC CIRCUIT

90. CONCEPT OF MAGNETO MOTIVE FORCE AND DEFINITION OF RELUCTANCE

THE rapid development of electromotive and electrogenerative machines beginning about 1870 made a development of methods of calculations and treatment of the various fields produced by electromagnetic devices imperative. A very convenient method of treatment was devised largely by H. A. Rowland, an American physicist, in 1873. He made use of the idea that a current flowing in a conductor produces a magnetic field, that is, a flux of magnetic lines of force through the circuit. Since to drive a unit magnet pole around a circuit against the field representing this flux requires work, Rowland in analogy to the work done by an electromotive force in driving current through a circuit called the work to carry the magnet pole once around the magnetic circuit, or system, the *magneto motive force.* Thus by dividing magneto motive force by flux in analogy to dividing electromotive force by current one gets a quantity Z called the reluctance of the magnetic circuit equivalent in its mathematical behavior to the quantity called resistance in the flow of electrical currents. Since means are devised to calculate the reluctance of the elements of a circuit, as is the case for resistance also, the reluctance of a circuit can be determined. This gives the flux if the magneto motive force is known, or if the flux is given one can calculate the magneto motive force needed to give that flux. It also happens that the reluctances in series are additive and when placed in parallel they follow the same laws as resistances in parallel. It is thus possible to compute the flux in branched circuits of all sorts. It must however be borne in mind that the analogy is purely a formal one, for in the case of currents there is actually a transfer of electricity while in magnetism there is no actual flow.

With this outline of the situation we can proceed to a quantitative formulation of the problem with illustrative applications. It was found in Chapter XVII that the work to carry a unit pole about

a coil of n turns carrying a current i_a was given by $W = 4\pi n i_a$ where i_a is in absolute units. If we speak of the current in amperes then $i_a = \frac{i}{10}$ and the work per unit pole becomes $W = \frac{4\pi n i}{10}$. Thus the work is proportional to ni, that is, to the ampere turns. This work per unit pole in analogy to work per unity quantity in the case of currents (electromotive force) is called the magneto motive force. Thus magneto motive force or $MMF = \frac{4\pi n i}{10} = \frac{4\pi}{10}$ ampere turns. As a result of EMF in a circuit a current i flows. In analogy, a magneto motive force produces a flux of φ lines in the magnetic circuit. For the case of the current $\frac{EMF}{i} = R$, the resistance, a constant of the materials, their distribution and dimensions in the circuit. In the case of the magneto motive force $\frac{MMF}{\varphi} = Z$, where Z is called the *reluctance* of the circuit and represents the equivalent of resistance (magnetic resistance) to the flux φ in a circuit. This is also a characteristic constant of the magnetic materials, their distribution and dimensions in the circuit considered.

It has become the custom in recent years to redefine magneto motive force as the actual ampere turns $MMF = ni$ and to write

$$\frac{\frac{4\pi}{10} MMF}{\varphi} = Z.$$

The $\frac{4\pi}{10}$ then comes into the equations as a constant divisor and we write

$$\frac{MMF}{\varphi} = \frac{Z}{\frac{4\pi}{10}} = \frac{10Z}{4\pi}.$$

91. THE VALUE OF THE RELUCTANCE

One may at once deduce the value of the reluctance for a circuit as follows. Consider a complete magnetic circuit as that shown in Fig. 89. We can trace the lines of force all around the circuit. The work done in threading once around the circuit for unit magnet pole is $\frac{4\pi n i}{10}$. This work is however the sum of the products Hdl all around

the circuit, i.e., force H on unit pole times distance. Here H is the value of the field in a small length of the circuit dl within which H is sensibly constant. Thus

$$W = \oint H dl = \frac{4\pi ni}{10},$$

where the sign $\oint$ means the integral, or sum, of the quantities Hdl around the circuit, for in general H varies along the circuit. This integral is called the *line integral of the field* and is one of great importance in all problems dealing with circuital relations of any sort, electrical or magnetic. For a circuital tube of induction (i.e., a tube of force for which the total number of lines is constant) we can write $\varphi = BA = \mu HA$, where B is the induction, A the area of the tube which may vary, and μ is the permeability.

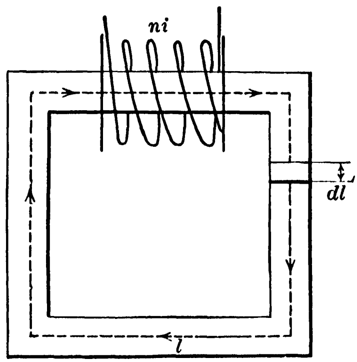


FIG. 89.

Thus $H = \dfrac{\varphi}{\mu A}$, and the equation above becomes

$$W = \frac{4\pi}{10} MMF = \oint H dl = \oint \frac{\varphi dl}{A\mu}; \text{ or } \frac{4\pi}{10} MMF = \varphi \oint \frac{dl}{\mu A}.$$

Thus

$$\frac{MMF}{\varphi} = \frac{10}{4\pi} Z = \frac{10}{4\pi} \oint \frac{dl}{\mu A}.$$

In general, A is a function of l and integration must be carried out over the circuit. Frequently, however, the circuit may be broken into parts in which A is constant over a length l and $Z = \Sigma \dfrac{l}{\mu A}$. This says that Z is the sum Σ of the terms $\dfrac{l}{\mu A}$ in the circuit, in which over a length l, A is constant so that $\displaystyle\int \frac{dl}{\mu A} = \frac{l}{\mu A}$. Thus we see that in a magnetic circuit it is frequently possible to express Z as the sum of a group of component reluctances $Z_1 = \dfrac{l_1}{\mu_1 A_1}$, $Z_2 = \dfrac{l_2}{\mu_2 A_2}$, $Z_3 = \dfrac{l_3}{\mu_3 A_3}$, and so forth. Since $Z = \dfrac{l}{\mu A}$ we see again a close analogy to resist-

ance, for in the case of resistance $R = R_0 \frac{l}{A}$, where R_0 is similar to $\frac{1}{\mu}$, and l and A are strictly analogous. We may also note that the reluctances Z in series are additive as are resistances R. For two reluctances in parallel we may again note that as for resistances $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$ so with reluctances $\frac{1}{Z} = \frac{1}{Z_1} + \frac{1}{Z_2}$. Thus we can easily compute the values of the reluctance for a magnetic shunt. In general, then, we may write

$$\varphi = \frac{4\pi}{10} \frac{MMF}{\Sigma \frac{l}{\mu A}} = \frac{4\pi}{10} \frac{\text{ampere turns}}{\Sigma \frac{l}{\mu A}}.$$

Applications.—Three types of problems may be considered. (1) One may attempt to calculate φ for a given circuit such as those depicted below in Figs. 90 and 91. (2) Again one may wish to obtain a given flux φ in a region, and having fixed on the form of circuit one must know how many ampere turns will be needed to give the flux. (3) Finally one may find it necessary to calculate the effect of a magnetic shunt.

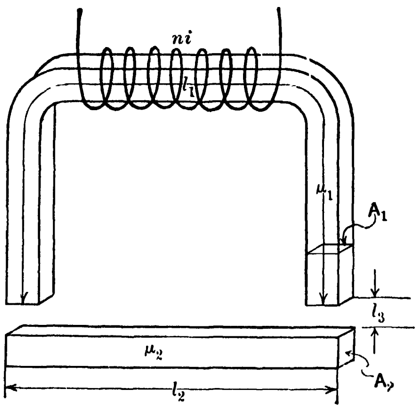


FIG. 90.—Calculation of Flux through Magnet and Keeper.

(1) Given the circuit of a horseshoe magnet as shown in Fig. 90, with the coil having an ampere turn value ni, with an area of cross section A_1, a permeability μ_1 and a length l_1 in the horseshoe part. The iron keeper separated from the magnet by a small air gap at the two poles of the horseshoe has a length l_2, an area A_2 and a permeability μ_2. The lengths l_3 of each of the two air gaps is so short that we may consider the area A_3 equal to A_1. For these air gaps $\mu = 1$, and we see that in this case most of the reluctance comes in the gap as μ_1 and μ_2 are in the thousands. The solution of the problem is then merely

$$\varphi = \frac{\frac{4\pi}{10} ni}{\frac{l_1}{\mu_1 A_1} + \frac{l_2}{\mu_2 A_2} + \frac{2l_3}{A_1}}.$$

(2) Given the same circuit as in the preceding problem. It is required to design a coil for the horseshoe magnet to give a flux φ in the iron keeper. The problem is now to be complicated by the fact that the keeper has been removed a greater distance from the horseshoe so that l_3 is greater, and by the fact that μ_1 and μ_2 are undefined. When l_3 is fairly large a certain percentage of the flux that previously went through the keeper prefers to go through the air from one pole to the other and not through the keeper. Thus there is a loss of f per cent of lines for the keeper. In such problems the types of iron used (i.e., μ_1 and μ_2) are probably fixed as are the dimensions A_1 and A_2, but μ_1 and μ_2 must be calculated for the value of B required to be used in the circuit. It is possible that the areas A_3 are somewhat increased, but for simplicity these may be considered the same. The problem thus becomes one of calculating the ampere turns ni or the M.M.F. to cause φ lines to pass through the keeper with a loss of f per cent of lines due to leakage. The ampere turns can best be calculated by finding the ampere turns $n_1 i_1$ to force this flux through the portion characterized by l_1, A_1, μ_1, so that φ lines go through $l_2A_2\mu_2$, the ampere turns n_3i_3 to force the flux φ through the air gap l_3A_3 and the ampere turns n_2i_2 to force the flux φ through the keeper $\mu_2l_2A_2$, and adding them. Since f per cent of the flux in $l_1A_1\mu_1$ is lost before it passes through $l_2A_2\mu_2$ there must be $\dfrac{100}{100-f}\varphi$ lines generated in $A_1l_1\mu_1$. Now if $\dfrac{100}{100-f}\varphi$ lines go through $A_1l_1\mu_1$ the value of the induction B_1 is $\dfrac{100}{100-f}\varphi/A_1$ for this part of the circuit. Again if φ lines are to thread $A_2l_2\mu_2$, the induction B_2 is $\dfrac{\varphi}{A_2}$. Hence we can at once determine μ for these values of B from tables characterizing the particular type of metal used. Thus as soon as the values of B_1 and B_2 are computed the values of μ_1 and μ_2 can be found. We then write since $ni = \dfrac{10}{4\pi}Z\varphi$

$$n_1i_1 = \frac{10}{4\pi}\frac{100}{100-f}\varphi\frac{l_1}{\mu_1A_1}$$

$$n_3i_3 = \frac{10}{4\pi}\varphi\frac{2l_3}{A_3}$$

$$n_2i_2 = \frac{10}{4\pi}\varphi\frac{l_2}{\mu_2A_2}.$$

As $ni = n_1 i_1 + n_2 i_2 + n_3 i_3$, we have

$$ni = \frac{10}{4\pi}\varphi\left(\frac{100}{100 - f}\frac{l_1}{\mu_1 A_1} + \frac{2l_3}{A_3} + \frac{l_2}{\mu_2 A_2}\right).$$

The value of ni being now calculated, the wire and dimensions of the coil must be designed to stand the current without overheating and yet give the adequate ampere turns.

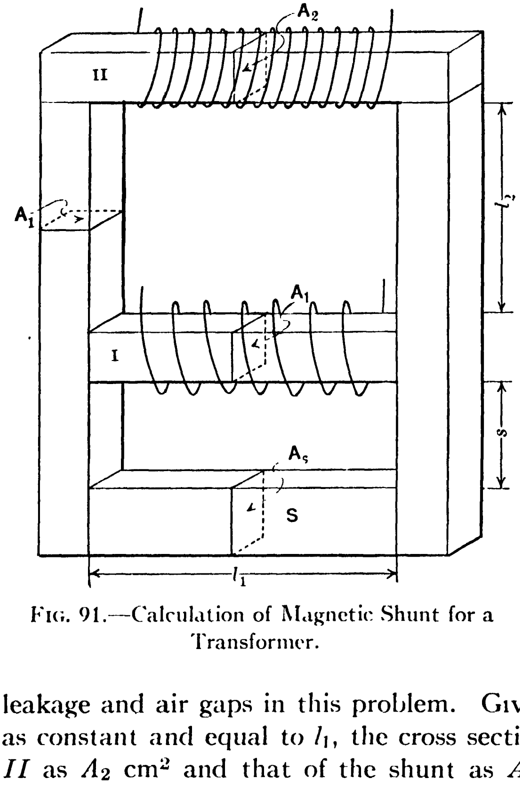


FIG. 91.—Calculation of Magnetic Shunt for a Transformer.

(3) Suppose one is given the circuit depicted in Fig. 91, which is the sort of a branched circuit sometimes met in laboratory transformers to enable an easy reduction in potential to be achieved. It is desired to calculate the distance at which the cross-arm shunt should be placed relative to the primary branch I so that by putting it into the circuit the flux through the secondary coil II can be reduced to one-half of its value in the absence of the shunt. One may neglect leakage and air gaps in this problem. Given the width of the frame as constant and equal to l_1, the cross section of I as A_1 cm^2, that of II as A_2 cm^2 and that of the shunt as A_S cm^2, while the distance between I and II is l_2 cm and that between I and S is S the unknown required. Now the total flux generated is $\varphi_0 = \dfrac{4\pi}{10}\dfrac{ni}{Z_0}$ when the shunt S is absent. Here $Z_0 = Z_I + Z_{II}$. When S is in place the total flux is $\varphi_S = \dfrac{4\pi}{10}\dfrac{ni}{Z_1}$, where Z_1 is given by $Z_1 = Z^1{}_I + \dfrac{Z^1{}_{II} Z_S}{Z^1{}_{II} + Z_S}$, which follows from the law for reluctances in series and parallel in analogy to resistances. Of this flux φ_S, the fraction $\dfrac{\varphi_{II}}{\varphi_S} = \dfrac{Z_S}{Z^1{}_{II} + Z_S}$ is driven through the secondary II. This follows in analogy to the

tact were perfect, there would be no stray flux. Between these two states, as the air gap is increased, all degrees of loss of flux or leakage can occur, as shown in Fig. 94*b*. These must be taken account of when one considers circuits, as was done above.

(3) **The Effect of Air Gaps.**—Besides increasing the leakage, the effect of an air gap on the hysteresis loop, or the $B - H$ curve, is to make higher fields essential to produce the same degree of magnetism. The $B - H$ curve will thus take on an elongated shape instead of the more compact shape. This is due to the fact that μ for air is small and to get the same flux we must increase the M.M.F.

(4) **The Structure of Electromagnets.**—Electromagnets are of two types. Where a high resistance is permissible and a small current flow is required, the magnet is wound with a large number of turns of as low a resistance as possible. Then for a small current, the magnetic effect is a maximum. In some cases where it is possible to get very high currents, as from some of the modern high current generators, the number of turns wound on a magnet are comparatively few but the currents are prodigious. To prevent heating with such currents, the armature windings are made of tubes through which water is circulating to carry off the excess heat. In general, these are wound about magnetic circuits with poles close together. Fields of the order of 50,000 gauss can be obtained with such magnets in short air gaps. Higher magnetic fields have been obtained by Kapitza for short intervals of time by the discharge of a high capacity storage cell of low internal resistance through a solenoid.

(5) **Lifting Power of Magnets.**—We finally come to the question of the lifting power of magnets. When we bring two magnets close together, the plane surface of the magnet, or magnet and keeper with induced magnetism, *act like magnetic shells* * with a density of magnetization per cm^2 equal to σ. The number of lines going out from one surface to the other is $2\pi\sigma$ per cm^2. The force on the density σ of one pole exerted by the $2\pi\sigma$ lines from the other pole is the product of those two and is expressed by $f = 2\pi\sigma^2$. If there is no divergence of lines of force, which is the case when the two surfaces are parallel

* A magnetic shell is merely a thin piece of magnetized material whose area has linear dimension of much greater magnitude than its thickness. If this shell is magnetized with all its north poles on one surface and the south poles on the other, it constitutes a magnetic shell. Unlike the case for a conducting metal surface with σ charges on it (whose *lines of force all emerge from the surface normal to it* as there is no field inside the conductor), the σ north (or south) poles on the surface of the shell send but one-half their lines of force outward into the space, the other half ($2\pi\sigma$) lines running from the north (south) to the south (north) pole through the shell as the shell is a magnetic conductor.

and close to each other, H (the field strength) equals $4\pi\sigma$ in the gap. Thus

$$\sigma = \frac{H}{4\pi}, \text{ or } f = \frac{H^2}{8\pi}.$$

Thus the force between the two surfaces of area 1 cm^2 is

$$f = \frac{H^2}{8\pi}.$$

If there is no divergence of the lines, that is, if there is no stray field, which is true for a short gap, $H = B$, where B is the induction through the iron, and we have

$$f = \frac{B^2}{8\pi}, \text{ as the force } per\ cm^2.$$

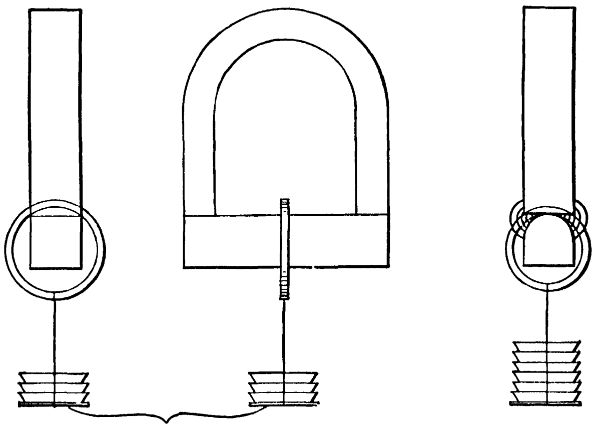


FIG. 95.—Experiment Showing the Lifting Power of a Magnet to be Proportional to AB^2.

Thus if we know the induction through the iron we can calculate the force of traction between magnet poles; or vice versa, if we measure the lifting power of a magnet, we can determine the induction through it provided the air gap is small. If the area of the poles is A the force f per cm^2 must be multiplied by A to give the total force f_A on area A. That is, the total force is $f_A = \dfrac{AB^2}{8\pi}$. The fact that the attractive force between a magnet and a piece of iron is proportional to the area times the square of the induction is nicely illustrated by the experi-

ment shown in Figs. 95*a* and 95*b*. In this experiment, which depicts the side view of an iron horsehoe magnet supporting a keeper carrying a scale pan with weights by means of string, we will observe that if the keeper is perfectly flat (Fig. 95*a*), it will support less load than the keeper with the curved surface in contact with the iron, illustrated in Fig. 95*b*. In the first case while the area A is large the induction is the total flux of the magnet through the keeper divided by the area of the keeper in contact with the magnet. In the second case, since the area A of the keeper in contact with the magnet is a small fraction of the area in contact in the first case, the induction, except for the small loss of lines of force due to leakage, is increased in inverse proportion to the area of contact; for in this case, except for the flux lost to leakage, most of the flux is found to cover a much smaller area and thus B is increased accordingly. Therefore, B^2A is distinctly greater in the second case than in the first, as shown by the greater traction.

CHAPTER XX

INDUCED ELECTRIC CURRENTS

93. DESCRIPTION OF INDUCED CURRENTS, LENZ'S LAW, DIRECTION OF INDUCED E.M.F.

In 1831 Faraday made a discovery the consequences of which have led to the development of modern industrial electrical engineering. As it had been observed that a current produced a magnetic field, he decided that vice versa a magnetic field should produce a current. In looking for this effect he found that in fact under certain conditions a magnetic field will produce a current in a conductor. From the first discovery of the effect, in ten days' time, he had completed an investigation of the phenomenon, and our present-day knowledge of the phenomenon hardly exceeds what he discovered in those ten days.

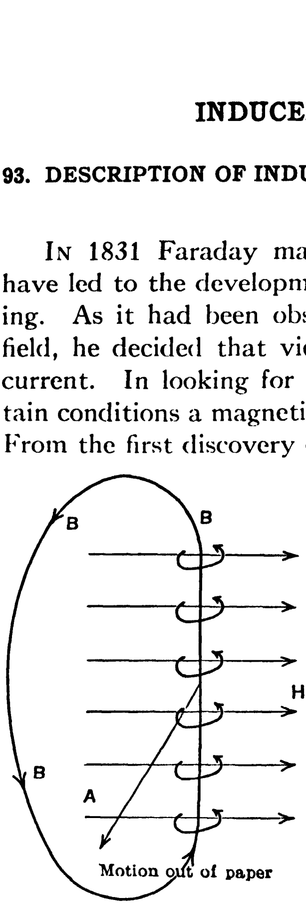


Fig. 96.—Direction of an Induced Current Caused by a Conductor Cutting Lines of Force.

Faraday's observation was this—that if a conductor be moved in a magnetic field in such a way as to have a component of its motion perpendicular to the field, an E.M.F. is induced at the ends of the conductor. If the conductor is short-circuited or the ends are joined, a current flows through the conductor. The direction in which this current flows can be deduced at once from the law of conservation of energy.

The law for the direction of flow of current is known as *Lenz's law.* The current set up in the conductor must flow in such a direction as to build up a magnetic field opposing the change in the circuit. Otherwise, the magnetic field set up by the current flowing would increase the motion and generate more current. This process could not go on indefinitely increasing the motion and building up more and more current at the expense of nothing. In a more general form,

Lenz's law could be stated by saying that the current induced by any magnetic changes in the regions around a conductor flows in such a way as to produce a magnetic field which opposes the change or results in motions such as to oppose the change. This is illustrated for one case in Fig. 96. The magnetic field H is given by the horizontal lines. The vertical wire is being moved outward from the paper as indicated by the arrow A. The electrical current generated flows around the circuit in the direction of the arrows B and goes upward in the conductor. The magnetic field produces currents in the wires so that the side of the wire outward from the paper has the lines of force flowing in the same direction as the applied external field H. From Chapter XVII it will be seen that this results in a force on the wire urging it into the paper.

94. MAGNITUDE OF INDUCED E.M.F.

The magnitude of the E.M.F. may be calculated as follows. The magnetic field H of Fig. 97 is represented by dots, the lines of force running out of the paper. Two conducting wires connected to a resistance coil form a sort of track. The distance between them is l. A conductor C is laid perpendicular to these and is moved a distance dx parallel to the rails. Suppose that moving the conductor generates an E.M.F. of value E. Suppose also that this E.M.F. causes a current i_a to flow through the wire during the time dt while the conductor is moved dx. The work done $(W) = Edq = Ei_a dt$. On the

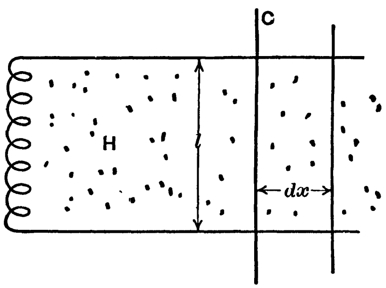


FIG. 97.—Calculation of the Magnitude of an Induced E.M.F.

other hand, if the current i_a is flowing through the conductor as a result of the induced E.M.F., we know that the force f, acting on i_a due to the field H which opposed the motion will be given by $f = i_a lH$, therefore the work done $(W) = fdx = i_a lHdx$. Thus, since this work is the same work as that done against the E.M.F., we have

$$Ei_a dt = i_a lHdx,$$

or

$$E = Hl\frac{dx}{dt}.$$

Now ldx is A, the area of the field swept out, but the area of the field swept out is nothing else than the change in area inside the circuit,

and multiplying this change in area by the number of lines of force per unit area we have the *change in flux* through the circuit. If we let $Hldx = d\varphi$, the change in flux, our equation becomes

$$E = \frac{d\varphi}{dt}.$$

that is, *the E.M.F. is rate of change of flux in the circuit.*

There is thus a current set up by the E.M.F. E caused by cutting of lines of force where $\frac{d\varphi}{dt}$ is the rate of cutting the lines of force.

95. DEFINITION OF E.M.U. OF POTENTIAL

We can then *define the electromagnetic unit of potential by saying that it is the potential difference produced by a given rate of cutting of lines of force. By this equation* $E = 1$ *E.M.U. of potential when one line of force is cut per second. The volt is the E.M.F. produced when* 10^8 *lines of force are cut per second.* This is a definition of potential on the electromagnetic system which is a perfectly legitimate one and could replace our fundamental work definition. Since, however, it is directly related to the work done by means of the work definition which we have already given, and inasmuch as it connects our unit to the absolute C.G.S. system, the original definition of the unit of potential difference on the electromagnetic system is the one which will be adhered to.

96. E.M.F. PRODUCED BY CUTTING LINES OF FORCE

It makes no difference whether the conductor moves and lines of force are cut or whether the lines of force move and cut the conductor. This fact has led to a loose definition of the process by which the E.M.F. is produced, for it has led to the statement that the E.M.F. is produced and is proportional to the rate of change of magnetic flux through the circuit. It can be shown that it is *not the rate of change of flux, but the cutting of the lines of force associated with the rate of change of flux, which causes the E.M.F.* Thus $\frac{d\varphi}{dt}$ is merely an expression for the rate of cutting of the lines of force. For example, it is possible to obtain an E.M.F. by cutting lines of force while the flux is not changed through the circuit. In Fig. 98, N is the north pole of a soft iron bar magnetized by means of the coil C at its lower

end. From this pole lines of force emerge and swing around downward to the south pole. W is an L-shaped wire loop whose lower end dips into a trough of mercury marked Hg, pivoted so that the loop can turn freely about N. Between N and the mercury trough there is connected a galvanometer G through a wire. It will be observed that as the wire W rotates around the magnet as an axis a current is generated in the circuit and is recorded by the galvanometer. It is obvious that since the field around the magnetic pole N is uniform, the flux through the wire W and through the circuit as a whole is constant. That is, the lines of force through the circuit do not change. The upper part of the wire W, however, is continually cutting the downward component of the lines of force emerging from N, with the result that an E.M.F. is caused while there is no change of flux in the circuit.

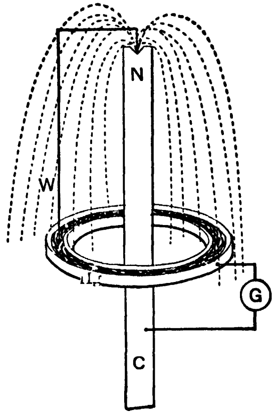


FIG. 98.—Experiment Showing the E.M.F. to be Generated by a Cutting of Lines of Force.

97. APPLICATION TO EARTH INDUCTOR

We now turn to an application of the principle of magnetic induction which enables us to measure the earth's magnetic field and compare it to our current measuring system based on the magnetometer. The instrument used enables us to determine the earth's field more easily and quickly than the absolute magnetometer method. Suppose we mounted a rectangular coil of n turns about a horizontal axis and suppose it placed so that its axis is perpendicular to the lines of force of the earth's field. Then, as it rotates about its axis, the components of the n wires of the rectangular coil parallel to the axis of the coil cut the earth's lines of force at right angles. An E.M.F. is then induced in the coil and may be measured as detailed below. This instrument is called the *earth inductor*. Assume an earth inductor having a coil of n turns, and that the earth's field is H_T. Assume that the area of the coils is A. As the earth inductor is rotated the areas perpendicular to the earth's field swept out in equal intervals of time dt by the component of the wires parallel to the axis of the inductor are not equal. Thus the E.M.F. generated is not constant, and we can write for the instantaneous value of the E.M.F. the expression

$$E_{\text{inst.}} = \frac{d\varphi}{dt} = H_T n \frac{dA}{dt}.$$

In the first quarter turn the two ends of the rectangular coil together cut an area of magnetic flux A equal to the area of the coil, that is, they cut $H_T A$ lines of force. In the second quarter turn, the lines of force cut are again $H_T A$ for each wire of the conductor, thus $2H_T A$ lines of force are cut in one-half turn. In the second half turn, lines will be cut with the coils moving in the opposite sense, and the current would flow in the opposite direction. The current can, however, be rectified by a commutator. We may consequently confine ourselves to the motion for one-half a turn only. The average E.M.F. for n wires for one-half turn is then

$$E = \frac{2nH_T A}{\tau}$$

where τ is the time of one-half revolution. Since $i_a = \frac{E}{R}$, then $i_a = \frac{2nH_T A}{R\tau}$.

But

$$i_a\tau = q = \frac{2nH_T A}{R} \text{ in absolute E.M.U.}$$

Now we can measure q by what is known as the ballistic galvanometer (see Chapter XXIII). This is an instrument in which the time of deflection is large compared to the time during which q passes. It is also an instrument in which there is little damping. In such an instrument

$$q = \frac{K'T\theta\rho^{1/2}}{2\pi}.$$

Here K' is the galvanometer constant, and θ is the deflection of the galvanometer when the field H_T is cut, ρ is the damping factor and T is the period of the galvanometer. Thus, knowing q, H_T can at once be obtained. This quantity H_T which is measured is the *total intensity* of the earth's magnetic field. The horizontal component H may then be obtained by knowing the angle of dip.

98. APPLICATION TO A SECONDARY COIL ABOUT A PRIMARY IN WHICH THE CURRENT CHANGES

Another case of interest is one where an E.M.F. is induced in one coil of wire by a change of current in a neighboring coil. This case is of considerable practical importance in that it is used in the measurement of magnetic induction and in estimating the values of mag-

etic field strengths. Assume that we have the long coil represented n Fig. 99 by n_1, and that it has n_1 turns of wire in it. Assume that ts area is A and that it has a current of i amperes flowing through it. bout it is a second small coil of n_2 turns. The ends of this small coil re connected to a galvanometer G of the allistic type. While the current i is flowing hrough the large coil n_1 there is a flux φ qual to

$$\frac{4\pi n_1 i A}{10 l}$$

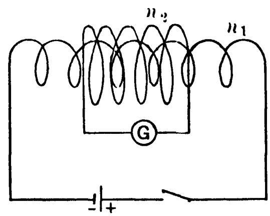


FIG. 99.—E.M.F. Induced in a Secondary Coil.

hrough it, where i is in amperes. If the urrent in n_1 be suddenly broken the magetic flux goes from φ to 0 in a time τ which is nknown. The time depends on the way in which a spark is drawn ut at break. Thus on breaking the circuit the average change in flux er unit time is

$$\frac{\varphi}{\tau} = \frac{4\pi n_1 i A}{10 l \tau}.$$

n one turn of the coil n_2 the average E.M.F. in volts will then be

$$E = \frac{4\pi n_1 i A}{10^8 \times 10 l \tau}.$$

or n_2 turns the average E.M.F. in volts will be

$$E = \frac{4\pi n_1 i A n_2}{10^9 \tau l}.$$

f τ is known, only the average value of E_2, the E.M.F. in the coil n_2, may be computed; for it is obvious that as the current i dies out on-uniformly, the rate of cutting of lines of force will not be uniform, nd thus E_2 will not be uniform. The study of i as a functon of time which would be needed in an accurate investigation of this circuit is nly possible through the use of the cathode ray oscillograph. For pracical purposes this may not be possible, and we must make use of the ame device mentioned in connection with the earth inductor, i.e., he ballistic galvanometer. Thus, if E_2 is short-circuited through a allistic galvanometer of resistance R the average current flowing s then given by

$$\bar{i}_2 = \frac{E}{R} = \frac{4\pi n_1 n_2 i A}{10^9 \tau l R}.$$

The quantity Q of electricity in coulombs which flows is therefore

$$Q = \bar{i}_2\tau = \frac{4\pi n_1 n_2 i A}{10^9 R l}.$$

The ballistic galvanometer when calibrated measures Q. Consequently we can measure the flux through the coil. As was stated before, if this flux be measured with iron in the coil and then in the absence of iron a simple calculation will give the change in flux produced by the iron and hence its permeability. If i is in amperes and R is in ohms then Q is measured in coulombs.

99. APPLICATION TO FARADAY'S DISC

The Faraday Disc.—Faraday was the first to show the continuous production of an E.M.F. by cutting a magnetic field. The square H in Fig. 100 represents a magnetic field perpendicular to the plane of a copper disc of radius l mounted on a horizontal axis C. This disc is rotated so that the angle $d\theta$ is swept out in the time dt. Its lower end dips in the mercury in the trough Hg. The axle and the mercury trough are connected to a galvanometer G. The E.M.F. E_{volts} generated will be then given by

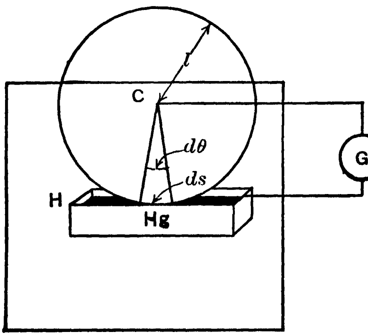


FIG. 100.—Faraday Disc.

$$E_{\text{volts}} = \frac{H}{10^8}\frac{dA}{dt},$$

where dA is the area swept out in dt. Now

$$\frac{dA}{dt} = \tfrac{1}{2}l\frac{ds}{dt},$$

where ds is the element of arc swept out in the time dt. But

$$\frac{ds}{dt} = \frac{l\,d\theta}{dt}.$$

Therefore

$$\frac{dA}{dt} = \frac{\tfrac{1}{2}l^2 d\theta}{dt}.$$

Now θ equals 2π radians for one revolution. For N revolutions, it equals $2\pi N$ radians. Thus

$$\frac{d\theta}{dt}$$

for N revolutions per second is equal to $2\pi N$. Whence

$$E_{\text{volts}} = \frac{H}{10^8}\,\tfrac{1}{2}l^2 2\pi N.$$

As

$$\pi l^2 = A \text{ (the area of the disc).}$$

$$E_{\text{volts}} = \frac{HAN}{10^8}.$$

CHAPTER XXI

SIMPLE DYNAMOS AND MOTORS

100. DERIVATION OF EQUATION FOR SIMPLE DYNAMO

IN the simple dynamo we have a rectangular coil of wire rotating about the axis A normal to the plane of Fig. 101; B and B' are the cross sections of the ends of the rectangle viewed perpendicular to the axis at two instants dt apart. The sides of the rectangle parallel to the paper are of no concern as they move parallel to the magnetic lines of force. The magnetic field in which the system moves is indicated by the horizontal lines H. The axis is perpendicular to these lines as is the vertical AC from which the angle θ of the armature with its neutral position C is measured. Assume that there are n_1 turns in the coil, that the length of the armature coil normal to the paper is l, and that the distance from the axis to the coil is r. Assume a field strength H. E_{dt}, the instantaneous electromotive force, is then $n_1 \frac{d\varphi}{dt}$, where $d\varphi$ is the change of flux through the coil during dt, and n_1 is the number of turns in the coil. Assume that the dynamo armature revolves through an arc ds in the time dt, when it makes an angle of θ with the neutral position. If it moves a distance ds, the distance moved perpendicular to the field is $ds \sin \theta$. Since the length of the conductor perpendicular to the field H is l, the area swept out in dt is $lds \sin \theta$.

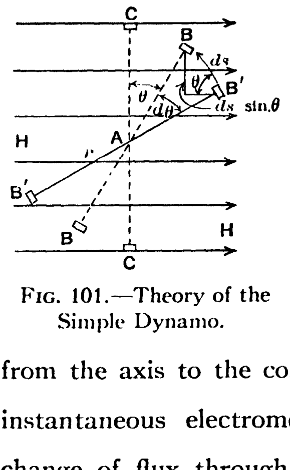


FIG. 101.—Theory of the Simple Dynamo.

The field has a strength H and there are n_1 turns in the armature, whence

$$n_1 \frac{d\varphi}{dt} = \frac{n_1 Hlds \sin \theta}{dt}.$$

Now $ds = rd\theta$, where r is the radius of the coil, and

$$E_{dt} = n_1 \frac{d\varphi}{dt} = 2rl \frac{n_1 H \sin \theta d\theta}{dt},$$

for the whole coil since the wires B' at the other end of the coil are also producing an equal E.M.F. in the same direction.

Again,

$$\frac{d\theta}{dt} = 2\pi N,$$

where N is the number of revolutions per second, so that E_{dt} in volts,

$$E_{vdt} = \frac{2rlNH2\pi n_1 \sin\theta}{10^8}.$$

Now $2rln_1 = A$ where A is the total area of the n_1 turns in the armature. Then

$$E_{vdt} = \frac{2\pi ANH \sin\theta}{10^8}.$$

It is thus seen that the E.M.F. generated by the dynamo depends on the angle θ, on the number of revolutions per second N, on the field strength H, and on the total area A of the armature.

101. DYNAMOS, ALTERNATING CURRENT, AND DIRECT CURRENT

We now turn to the question of dynamos and generators. It was found above that the E.M.F. of a simple dynamo was given by the expression

$$E_{vdt} = \frac{2\pi n_1 NHA_1}{10^8}\sin\theta,$$

where N was the number of revolutions of the dynamo per second, n_1 the number of windings in the armature, H the magnetic field strength, and A_1 the area of the armature. It is seen that the E.M.F. varies as the sine of the angle. If the E.M.F. generated by the coil were picked up by two brushes on the armature mounting making sliding contact with two rings connected to the two ends of the armature, one would actually get an electromotive force which varies sinusoidally with the angle of the armature and consequently with the time, as seen in Fig. 102A. Dynamos of this sort are used to a very great extent in generating power for transmission. Power of this sort can be generated at high potential and transmitted to stations where

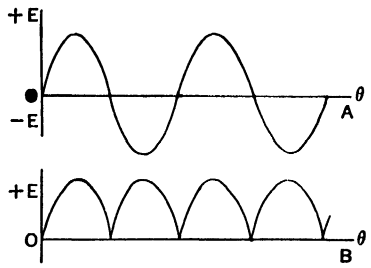


FIG. 102.—E.M.F. of a Simple Dynamo. A, From Collector Rings. B, Rectified by a Split Ring Commutator.

the potential is lowered by transformers to enable it to be used for lighting circuits. The equation below may be simplified to a more general form for the average E.M.F. given by a generator by replacing the quantity $2\pi IIA_1$ by a quantity U which is the average number of lines of force cut by a single wire of the armature in one revolution multiplied by 2π. The expression then becomes

$$E_{vdt} = \frac{n_1 N U}{10^8} \sin \theta,$$

and the maximum E.M.F. becomes

$$E_{\max} = \frac{n_1 N U}{10^8},$$

where $E_{\max}$ is the value of E_{vdt} when $\sin \theta = 1$. Use will be made of this expression later on.

In order to generate a *direct* current trom a dynamo use must be made of the *commutator*. This consists, in the case of the simple dynamo above, of a ring cut in two equal segments by saw cuts parallel to the axis at the ends of a diameter of the ring. The one segment is insulated by a small gap from the other segment. One end of the armature winding is attached to one segment; the other is attached to the other segment. Two fixed brushes at opposite ends of a diameter attached to the armature mounting make contacts with the ring so that as the armature turns through the position where it starts to cut the lines of force in such a way so as to reverse the flow of current, and thus reverse the potential, the commutator brushes make contact with the opposing segment. Thus one brush continually picks off only the positive side of the sine curve and the other brush picks off only the negative side. Since, however, there are positions where $\sin \theta$ is 0 it is seen that there are points in the revolution at which the potential difference between these two brushes is 0. If we again plot the potential difference against the time we will find a variation as indicated in Fig. 102*B* instead of the sinusoidal variation shown in Fig. 102*A*.

Such an electromotive force would give a fluctuating current through the circuit. To overcome this we multiply the number of coils in the armature. By multiplying the number of coils, which also means multiplying the number of segments in the commutator, the commutator segments are continually picking off the electromotive force of the coil that is passing through its maximum E.M.F.

This multiplication of armature coils results in smoothing out the

electromotive force given during one revolution, with the result that one has, instead of the sinusoidal E.M.F. given by an alternating current dynamo, an E.M.F. of the form shown in Fig. 103. The magnetic field of a dynamo may be excited by the dynamo itself or it may be excited separately as occasion demands. Self-exciting dynamos depend on their residual magnetism for starting the excitation. There is also a class of dynamos so wound that the armature wires do not cut the field in the same fashion as for the simple dynamo above. The wires are wound around an iron core as shown in Fig. 104. When the axis of the core is parallel to the field H, the flux is a maximum. When it is at right angles to the field, the flux is 0. Thus the current

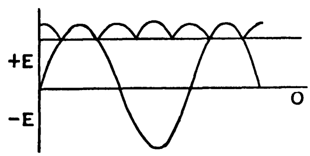


FIG. 103.—Rectified Alternating Current Using a Split Ring Commutator and Two Coils at Right Angles.

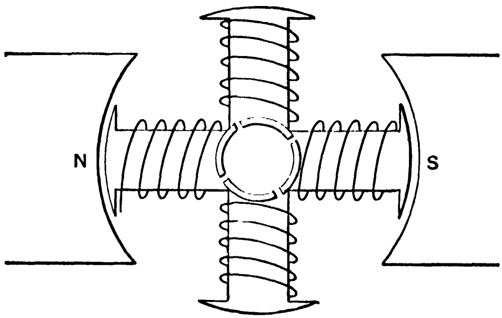


FIG. 104.—Alternative Method of Winding an Armature Coil.

is produced by a changing flux through the armature windings produced in a slightly different manner. This principle is also used in some high frequency generators in a somewhat different manner.

102. E.M.F. AND P.D. FROM A DYNAMO, EFFICIENCY OF A DYNAMO

If we generate an E.M.F. of E volts the current which comes out of the dynamo in accord with the same phenomenon discussed at the end of Chapter VII will be given by

$$i = \frac{E}{R_e + R_i}.$$

Here R_e is the external resistance of the circuit and R_i is the internal resistance of the windings. Again as in the case of batteries the potential difference V given across a certain external resistance R_e is related to the electromotive force by the relation

$$\frac{V}{R_e} = i = \frac{E}{R_e + R_i}.$$

If R_e is very large compared to R_t the V approaches E. For an open circuit ($R_e = \infty$), the V is equal to E. Furthermore, consider the expression

$$E = R_e i + R_t i,$$

and if V is equal to $R_e i$, therefore

$$V = E - R_t i.$$

It is thus seen that the potential V given depends on the E.M.F. E and on the internal resistance of the dynamo.

The power output of a dynamo is measured in watts. That is, it is measured in joules of energy per second. As will be remembered, the watt is the ampere multiplied by the volt, and the output of a dynamo would be V in volts times i in amperes. This is known as *the load on the dynamo.* The ideal efficiency of a dynamo is given by the expression

$$\text{Eff.} = \frac{\text{output in joules}}{\text{input in joules of mechanical work}}$$

$$= \frac{\text{output in watts}}{\text{mechanical input in joules per second}}$$

Ideally then as Vi is the output, and as all the mechanical work put in should give rise to electrical energy, one has input given by Ei. Hence

$$\text{Eff.} = \frac{Vi}{Ei} = \frac{Vi}{Vi + R_t i^2}.$$

Thus the smaller the $R_t i^2$ term, the more efficient the dynamo. As the current is determined by practical requirements, R_i should be as low as possible. Practically one cannot write input $= Ei$, as there are other losses of energy besides the i^2R loss in the armature.

The difference between the output of a dynamo and the mechanical work performed on it is caused by certain losses. These losses are as follows:

(*a*) Loss in iron core of armature due to eddy currents and hysteresis.

(*b*) Loss due to the iR drop in the field excitation due to R, the resistance of the magnet coils.

(*c*) The loss in the copper wire of the armature due to the i^2R_t heat production in the armature resulting from the flow of current against the internal resistance. If a heavy current is being delivered this loss is considerable. It is endeavored to make the internal resistance of the armature winding as low as possible.

(*d*) The loss in friction at the bearings of the armature and at the brushes. These losses all appear as heat, and it will be observed that operating generators are warmer than the room. In fact, it may be noted that most good generators are rated for a given temperature at which they should normally run.

103. RECIPROCAL RELATIONS BETWEEN DYNAMOS AND MOTORS *

We now turn to the reciprocal relationship of the motor and dynamo. If we put a source of electrical potential across a dynamo, the current flowing through the armature of the dynamo causes the armature to rotate. The dynamo thus acts as a motor. On the other hand if we turn the armature of a motor by mechanical means and excite the field coils the lines of force are cut and an E.M.F. is generated. Whether the dynamo is turned by mechanical means or is driven as a motor, the generated E.M.F. is picked up by the commutator. In general, a motor run as a dynamo is not particularly efficient and the same can be said of the reverse. The reason for this lies in the conditions underlying the usage of the dynamo and motor. The dynamo requires, as was previously seen, that the internal resistance R_i be as small as possible in order that the potential V maintained be as great as possible for the external resistance applied. In the case of the motor it is essential that the resistance be comparatively high in order to reduce the current flow and the loss of energy in the armature windings.

If now we place an external potential V across a motor the force acting on the armature causes it to rotate, and as the force produced is constant the motor armature accelerates. If another action, which we will discuss presently, did not occur, acceleration would continue until the armature flew to pieces or until friction on the bearings equaled the force acting on the armature. However, since a rotating armature in the magnetic field of the motor produces an electromotive force, and since by Lenz's law the electromotive force opposes the applied potential causing the motion, it is to be expected that the electromotive force so produced will change the net potential and therefore cause the motor to cease accelerating. Thus when the motor begins to rotate it generates a *back electromotive force* E' which causes a current to flow in a sense opposite to that of the imposed

* As this course is intended to familiarize the engineering student with the principles of electricity, it has been considered expedient to discuss only those applications to engineering which clearly illustrate the fundamental principles involved. As a result, this discussion will be confined to only the simple case of motors with separate field excitation.

potential difference. Since, under these conditions, the only source of potential in the motor other than the back electromotive force is the iR drop $R_i i$ in the armature windings, we can write the equation of the motor in the simple form,

$$V - E' = R_i i,$$

where V is the applied potential and E' is the back E.M.F. This is merely an application of the second of Kirchhoff's laws to the armature of a motor. We can consequently write that

$$\frac{V - E'}{R_i} = i,$$

and that

$$V = R_i i + E'.$$

104. POWER CONSUMPTION, EFFICIENCY, AND TORQUE FOR MOTORS

If we multiply both sides of the equation by i we have

$$iV = R_i i^2 + E'i.$$

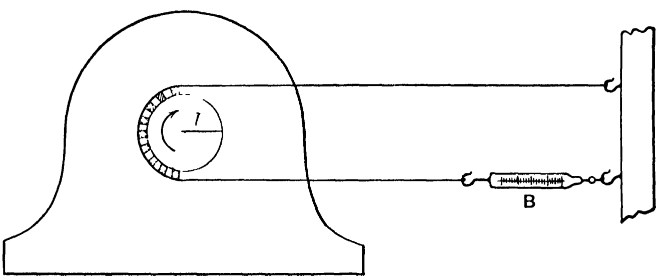


FIG. 105.—The Prony Brake.

The first term iV is the power which is put into the motor. The second term $R_i i^2$ is the power loss in the armature of the motor which goes to heating the motor, if we exclude eddy currents and other losses. The term $E'i$ is the power consumed in forcing the current i against the potential E'. This is the electrical power consumed in running the motor. Thus

$$E'i = \frac{\text{work}}{T} = \frac{W}{T} = \text{power},$$

where T is the time in which the work W is performed. Since work is equal to force times distance, $W = fs$.

$$\frac{W}{T} = \frac{fs}{T}.$$

Suppose the motor is turning over and working against the load on a prony brake. One form of the prony brake is shown in Fig. 105. It will be remembered that the brake is merely a belt placed against the circumference of a pulley of radius l attached to the motor. The force of the pulley on the belt can be measured by means of a balance B attached to the belt. The power output

$$\frac{fs}{T} = \frac{f2\pi lC}{T},$$

where C is the number of revolutions made in the time T and $2\pi l$ is the distance covered in a revolution, i.e., the distance over which f acted in one revolution. f in this equation is the force exerted by the prony brake, and is the one indicated by the balance. Therefore,

$$E'i = \frac{2\pi lfC}{T} = 2\pi flN,$$

Where N is the number of revolutions per second.

$$N = \frac{C}{T}.$$

The quantity fl is the force moment or torque G on the motor; that is, $fl = G$. As was seen in the generalized equation for the dynamo

$$E_{Vdt} = \frac{n_1 NU}{10^8} \sin \theta,$$

where U was the flux cut by each wire of the conductor in one revolution. Since for a D.C. motor, which we are now considering for simplicity, the E.M.F. is at its peak value, the expression for E_{Vdt} becomes $E_{max} = \frac{n_1 NU}{10^8}$, and this must be the back E.M.F. E' of our motor. Thus

$$E'i = \frac{n_1 NUi}{10^8} = 2\pi GN,$$

and the torque

$$G = \frac{in_1 U}{2\pi 10^8}.$$

We also can write that

$$Vi = R_i i^2 + \frac{n_1 NUi}{10^8},$$

and

$$V = R_i i + \frac{N n_1 U}{10^8}.$$

From the equation for G and that for V, we can draw the following conclusions: It is seen that as G gets less i, the current passing through the motor is less. Therefore, *for a given potential*, V, applied to the motor, with a decreased current i flowing through it, we have a corresponding increase in the term

$$E' \text{ or } \frac{n_1 N U}{10^8}.$$

Thus, the less the load the greater E' and the smaller the current i. That is, as we decrease the load we get a decrease in current through the motor, an increase in speed (for increase in

$$\frac{N n_1 U}{10^8}$$

can only mean an increased N), and a decrease in the $R_i i^2$ term. Since the ideal efficiency of the motor (in the absence of eddy current and hysteresis loss) is

$$\text{Eff.} = \frac{E'i}{Vi} = \frac{Vi - R_i i^2}{Vi},$$

we see that with a lighter load the efficiency of the motor is increased.

In designing a motor it must be so designed as to reduce the $R_i i^2$ loss and at the same time to get as high a back E.M.F., E', as possible. In general, the practice is to make n_1 high, to produce a powerful field, U, and to run the motor at a high speed. The large number of turns n_1 make the resistance of the motor fairly high. The limitation on increasing the efficiency of motors by extending the quantities in the direction indicated above is set by the limit on the speed of the motor required for practical purposes and the increase of R_i with too many turns. High speed motors require gearing down for practical use.

Gearing down has until recent years been a difficult problem in engineering practice. The necessity of proper gearing down of high speed rotors was the result of the development of the high-pressure steam turbine. Today, the rapid development in this field has led to the electrical drive in certain types of steam vessels. Thus, on ferry-boats and on certain battle-ships, the electrical drive is being used because of its great flexibility. The electrical drive is especially

important in the application of Diesel motors to ships inasmuch as these motors are not capable of flexible maneuvering. The gearing down, however, has the disadvantage of increasing frictional losses as well as wear and tear on the moving parts which can break down.

105. STARTING BOXES

It is obvious that if we suddenly close a switch, putting a potential V on a motor, a current

$$i = \frac{V}{R_i}$$

flows through the armature windings. As the resistance is after all comparatively low the high voltage which is applied to the low resistance R_i before the motor begins to turn over may give a current great enough to melt the wires of the armature. When running at full speed the potential forcing current through the motor is $V - E'$. The current i is thus decreased, the i^2R_i heating is then very much less and the motor operates safely. All heavy motors which are slow in gathering speed are therefore provided with a starting box. Such a starting box consists merely of a series of resistances which limit the maximum current i flowing through the motor armature. As the speed increases, each plug of the starting box gives a progressively lower series resistance and the motor is enabled to pick up speed without burning itself out.

CHAPTER XXII

INDUCTION

106. INDUCTION AND MUTUAL INDUCTION

WE now turn to a new concept in electricity and the last of our derived units. In 1831, Faraday found that when magnetic lines of force are cut by a conductor, an E.M.F. is generated. Lenz's law stated that an induced E.M.F. must be of such a sign that the current which it causes to flow will flow so as to make a magnetic field which maintains the previous condition of the circuit. That is, if we let lines of force due to the disappearance of a magnetic field cut a conductor the current will flow in the conductor in such a sense as to try to maintain the magnetic field which is disappearing. Again, if we make a magnetic field in a region where there was no field before in such a manner as to cut an electrical wire a current will flow in the wire in such a sense as to make a magnetic field tending to annihilate the field whose creation is causing the current. Thus, if we have two neighboring circuits, one of them in which a current can be made or destroyed, there will be produced in the other a current making a magnetic field in such a sense as to oppose any change in the initial conditions. Such a type of electromagnetic induction is called *mutual induction.*

We can compute the value roughly in the following way. Consider a coil 1 with a current i_1 flowing in it and assume that there are n_1 turns in that coil. Assume that it is near another coil and that to insure the same flux through both coils they be wound around a common iron core. This second coil 2 has n_2 turns in it. The flux in the iron due to the current i_1 in amperes will be

$$\frac{4\pi n_1 i_1}{10Z},$$

where Z is the reluctance of the magnetic circuit. If the current i_1 be suddenly stopped the average rate of a change of flux through both circuits is given by the expression

$$\frac{d\varphi}{dt} = \frac{4\pi n_1 i_1}{10Z\tau},$$

where τ is the time taken for the current to die out. In each of the n_2 turns of the coil 2 an E.M.F. will be induced equal to

$$\frac{d\varphi}{10^8\tau} \text{ volts.}$$

The average E.M.F. in volts in coil 2 is therefore

$$E_2 = \frac{4\pi}{10^9} \frac{n_1 n_2 i_1}{Z\tau}.$$

This will cause a current

$$i_2 = \frac{E_2}{R}$$

where R is the resistance of the circuit connected to 2. This current is the average current which will be generated while the current i_1 changes from i_1 to 0 in time τ. Since this current is changing in value continuously we generally find it more accurate to speak of the instantaneous electromotive force

$$E_{\text{Inst.}} = \frac{4\pi}{10^9} \frac{n_1 n_2}{Z} \frac{di_1}{dt},$$

and likewise, the instantaneous current i_2 will be given by

$$\frac{E_{\text{Inst}}}{R}.$$

In both the expressions there is a variable term $\frac{di_1}{dt}$ which depends on the rate at which the current in 1 dies out. There is also a constant factor

$$\frac{4\pi}{10^9} \frac{n_1 n_2}{Z}$$

which is a characteristic of the two circuits involved. It is seen that the instantaneous E.M.F. therefore is made up of *two* factors: the rate of change of current and a common factor characteristic of the shape of the circuit.

The ratio of

$$\frac{E_{\text{Inst.}}}{\frac{di_1}{dt}} = \frac{4\pi}{10^9} \frac{n_1 n_2}{Z}$$

is called the coefficient of mutual induction and is designated by the symbol M. *M, the coefficient of mutual induction, is thus a new constant characteristic of the two circuits.* We have now introduced the notion

of a new derived unit; namely, the unit which is defined as the ratio between the electromotive force produced and the rate of change of current in the activating circuit. It is seen that this is a derived unit involved in the case of transient currents only, and it depends, as is the case with all our derived units, on the dimensions of the circuit.

107. SELF-INDUCTION DEFINED

In the light of this we may now consider a still simpler case. A circuit has a current i flowing in it. Assume that i begins to die out. The magnetic field produced by the current i is now collapsing and the circuit is being cut by *its own lines of force.* By Lenz's law, this generates an E.M.F. tending to keep the current flowing. Thus, as the current is broken an E.M.F. is set up tending to maintain the current in its former state or vice versa. If we create a current in a wire an E.M.F. is set up opposing the flow of the current. That is, *every circuit has in it an electrical inertia tending to oppose any change in the electrical state of that circuit.* Thus, every circuit has an *inertia* which causes it to oppose any change in the state of the current within the circuit. This type of induction is known as *self-induction* because a change in the current induces a current in the circuit itself. Call φ the magnetic flux for *a unit current* in a given circuit. As the current i in the circuit changes, the flux φ is being changed. Therefore an E.M.F. of self-induction $E = \frac{\varphi di}{dt}$ is set up. *The flux for unit current change is the self-induction.* Symbolically,

$$\frac{E}{\frac{di}{dt}} = \varphi = L, \text{ the self-induction.}$$

108. UNITS OF SELF-INDUCTION

The absolute unit of self-induction on the electromagnetic system would be the self-induction which causes one absolute electromagnetic unit of E.M.F. to be induced in a circuit, when there is a rate of change of current in the circuit of one absolute electromagnetic unit per second.

The practical unit of self-induction is the value of the self-induction when one volt of potential difference is produced while the current is changing at the rate of one ampere per second. That is, L equals one practical unit when E equals one volt, and $\frac{di}{dt} =$ one ampere per sec-

ond. This unit is called the *henry* in honor of Joseph Henry, the American physicist, who first studied the properties of self-induction.

Since the volt is 10^8 absolute electromagnetic units of potential and the current in amperes is 0.1 of an absolute electromagnetic unit of current, the practical unit, the henry, equals $\frac{10^8}{0.1} = 10^9$ absolute electromagnetic units of induction. Thus the henry is *a very large unit.* The question next arises as to what the dimensions of self-induction are. Electromotive force, E, is work per unit quantity of electricity. Hence

$$E = \frac{W}{Q} = M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2} \text{ in the electromagnetic system.}$$

Current i has the dimensions $M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}$. The quantity

$$\frac{di}{dt} = \frac{i}{T} = M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-2}.$$

Self induction L has therefore the dimensions

$$\frac{E}{\frac{di}{dt}} = \frac{M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-2}}{M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-2}} = L, \text{ in the electromagnetic system.}$$

The electromagnetic unit of self-induction is therefore the centimeter, and the henry is 10^9 cm. The electrostatic unit of potential is 3×10^{10} electromagnetic units of potential, and since the electrostatic unit of current is

$$\frac{1}{3 \times 10^{10}}$$

electromagnetic units of current, therefore electrostatic unit of self-induction is 9×10^{20} absolute electromagnetic units of self-induction, or 9×10^{20} cm.

It is thus seen that the electromagnetic unit of self-induction is the smallest of the units of self-induction. It is this unit which has the absolute C.G.S. unit of length as its value. Thus, the electromagnetic unit of self-induction has the value of a cm, the henry is 10^9 cm and the absolute electrostatic unit of self-induction is 9×10^{11} henrys.

A milli-henry is 10^{-3} henry, and a micro-henry is 10^{-6} henry, hence a micro-henry is 1000 cm.

109. NATURE OF SELF-INDUCTION

Since L depends on the magnetic flux through the circuit which disappears when the current is changed, self-induction must depend very much on the form of the circuit. In the circuit of Fig. 106A, it will be seen that since the area enclosed in the circuit is practically 0, the self-induction will be practically 0. In the circuit of Fig. 106B, the self-induction will have a finite value depending on the area and the shape of the coil. The form of winding which doubles the wire back on itself allowing it to encompass a minimum of area, because of its low value of self-induction, is called a *non-inductive winding.* These windings are used in the making of resistance coils for resistance boxes. The value of L may be computed for certain very simple circuits. This has been done for coils of various shapes and sizes, and the information concerning such coils can be found in the bulletins of the Bureau of Standards. (Publications of U. S. Bureau of Standards, Scientific Papers, S 169, Dec. 18, 1916.) An analysis of this computation, which in most cases is exceedingly complex, must be left for more advanced courses. Of course, the presence of ferromagnetic substances increases L very much.

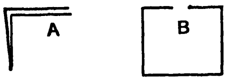


FIG. 106. — Effect of Shape on the Value of Self-Induction.

From what has been said about the value of the mutual induction we can set up the approximate expression for the value of the self-induction in the case of a circular coil of wire wound in such a manner as to have a calculable reluctance. The accuracy is limited by the closeness of the winding and the relation of the diameter to the length of the coil (i.e., end effect corrections). For a conductor of n turns having a current which changes by di in a time dt we have

$$\frac{d\varphi}{dt} = \frac{4\pi}{10}\frac{n}{Z}\frac{di}{dt},$$

As each of the n turns is cut by $d\varphi$ when the current changes di in a time dt, the E.M.F. E' of self-induction in volts is given by

$$E' = \frac{4\pi}{10^9}\frac{n^2}{Z}\frac{di}{dt} = L\frac{di}{dt}$$

whence

$$L = \frac{4\pi}{10^9}\frac{n^2}{Z}.$$

The units will be centimeters since Z is expressed in terms of electromagnetic units.

The self-induction is of prime importance in the study of transient or alternating current phenomena. In a sense, it acts like an inertia and as an inertia it will lead to electrical oscillations when introduced into the proper circuits. The importance of self-induction will be seen when the chapter on alternating current phenomena is reached. Needless to say, it is not important where we are dealing with steady or direct currents *except during the short interval while the current is being made or broken.* A striking example of the E.M.F. of self-induction is shown in the apparatus depicted in Fig. 107. In this a coil L with an iron core having a coefficient of self-induction of about 1 henry is connected to a switch S by which a direct current can be sent through L. Across L and in parallel with it is E, a small electric lamp. When S is closed with the steady current flowing, E glows dimly; on breaking the circuit by opening S, E will be observed to flash up brightly an instant before it goes out. This is due to the heavy current sent through E due to the self-induction of L on breaking the current at S.

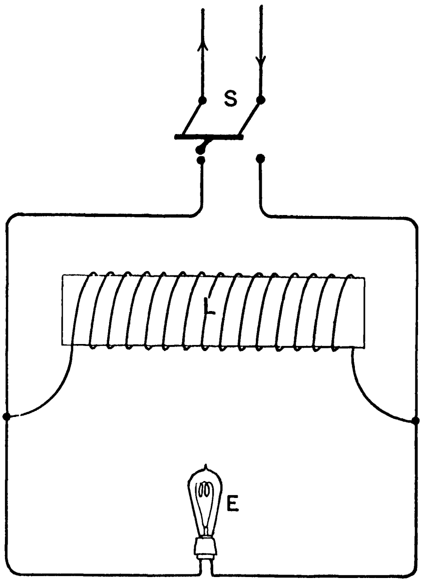


FIG. 107.—Illustrating the Effect of Self-Induction L on a Lamp E.

110. SELF-INDUCTION AND ENERGY LOSSES, EDDY CURRENTS

Induction and self-induction effects lead to many phenomena. One of these is of great importance because it causes energy losses. If we have a continuous metal body and place it in the neighborhood of a conductor with an alternating electrical field, induced currents are set up in the solid conductor which passing through the resistance cause heat production in the solid body. They are termed *eddy currents.* The heat production caused may be very intense and results in a loss of energy to the circuit acting on it. In fact, these induction effects are used frequently to heat bodies to exceedingly high temperatures in vacuo, and they are known as induction furnaces. In

such furnaces, high potential, high frequency electrical oscillations are set up in a coil surrounding the body to be heated. The eddy currents in this body will raise it to enormous temperatures. An excellent example of the effect of eddy currents is illustrated by the apparatus shown in Fig. 108. In this a pendulum, having a replaceable disc A which permits it to swing through a magnetic field H produced by a powerful electromagnet perpendicular to the plane of the drawing, oscillates for a considerable period in the absence of the field. The energy of the motion is then damped only by the friction in the bearing P. When the field is put on, the cutting of the magnetic force lines by the disc A as it swings through the field generates currents in A which are short-circuited in the highly conducting material. The energy of motion is thus completely consumed as A passes through

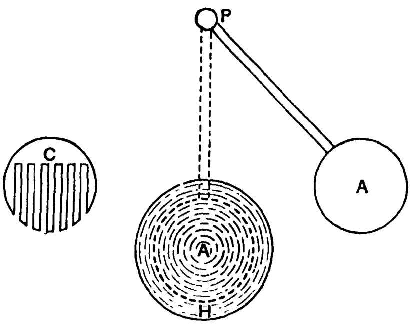


FIG. 108.—Damping Effect of Eddy Currents.

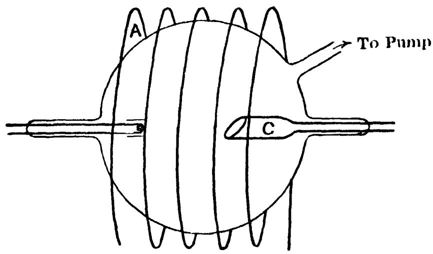


FIG. 109.—Use of the Induction Furnace in Outgassing an X-ray Tube.

the field H, and the disc acts as if it were cutting through a viscous liquid. If the disc A be replaced by the slotted disc C which reduces the scope of the eddy currents, the damping is much decreased and C will execute several oscillations before coming to rest. In Fig. 109, there is depicted an exhausted x-ray tube, fastened to a pump, having a metal anticathode C which is to be outgassed and heated to incandescence. About the tube is wound a coil A which is connected to a high frequency source of oscillations having considerable energy. The high frequency oscillations cause a rapid change in flux through the metal C, thereby generating numerous eddy currents which being short-circuited rapidly raise the metal to incandescence.

All iron cores in motor armatures and in transformers are subject to such losses as they permit the existence of small local eddy currents of self-induction. The losses may become quite great and reduce the efficiency of the transformer. It will be seen that the cores of trans-

formers and armatures of most motors are made up of groups of wires, or thin laminae of iron, separated from each other by shellac or some insulating body in such a way that the main induced currents in the wire cannot be set up, at the same time permitting the lines of magnetic flux to pass through the iron without interruption. In Fig. 110*A* is shown the frame of a transformer, while in Fig. 110*B* the end on view of the same transformer shows the laminated structure of the iron core *C*. It is seen that the eddy currents set up by the field in *I* would be such as to cause eddy currents to flow in the sense of the circular arrows *E* in *C*. The insulated laminations prevent this. The side view of the same core indicates that at the same time the

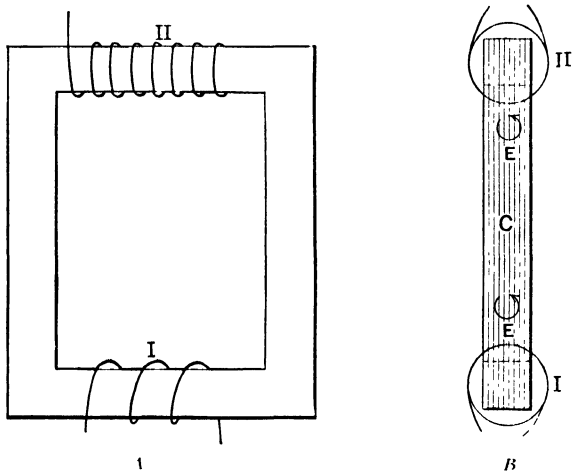


FIG. 110.—Illustrating the Laminated Structure of a Transformer Core.

continuity of each iron lamina insures a maximum of magnetic conductivity.

Another case of the manifestation of these eddy currents is in the damping produced by them. If one pass an instantaneous current through a ballistic galvanometer, then upon opening the circuit the galvanometer coil will oscillate back and forth until the friction in its suspension exhausts the energy applied. This may take hours. If, however, the galvanometer be shunted by a resistance the oscillations rapidly cease. The reason for this is obvious. The coil oscillating in the magnetic field sets up induced currents which in the coil and resistance are changed to heat. The heat energy thus generated uses the energy of vibration of the galvanometer in the production of these electrical heating effects, and the oscillations cease.

111. THE INDUCTION COIL

Another use of the property of induction is in the induction coil. In this instrument we have a primary coil I connected to a fairly high current source of potential. This primary circuit I, Fig. 111, is interrupted by some automatic breaking device B such that the current can be periodically made and broken. About the iron core C on which the primary is wound there are many turns of wire n_2 in another coil II called the secondary. The change of current and consequently of flux in the primary circuit results in the cutting of the secondary circuit by the magnetic field and causes multiplication of the number of turns cut by the magnetic flux to a high degree. On the basis of the analysis of mutual induction which has gone before, if we endow this secondary coil II with many turns the electromotive force E_2 set up in it is equal to

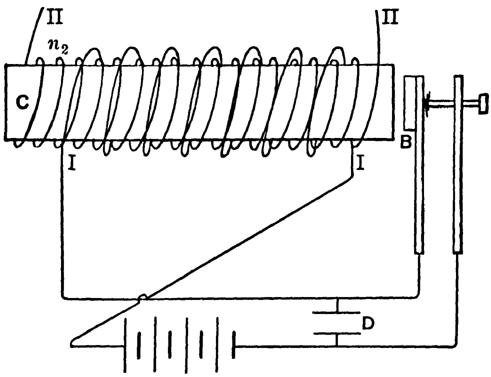


FIG. 111.—Diagram of an Induction Coil.

$$\frac{4\pi}{10^9}\frac{n_1 n_2}{Z}\frac{di_1}{dt}\text{ volts.}$$

If we use a small n_1 the potential necessary to drive a current i_1 through it need not be very large. On the other hand, the potential E_2 coming from the many turns n_2 of the secondary will be very high. It is, of course, necessary that Z be small in order to increase the magnetic flux. If n_1 is very great the self-induction of the primary is going to retard

$$\frac{di_1}{dt},$$

and consequently reduce the value of E_2. Thus, in the induction coil, we have a small number of turns n_1 and a high current in the primary, a fairly high magnetic flux, and a large number of turns n_2 in the secondary. The primary current is very much more effective on the *breaking* of the circuit than on the making of the circuit, for in the breaking of the circuit, the collapse of the field is very rapid, while the growth of current is retarded by the *self-induction* of I and the

mutual induction of *I* and *II*. The only factor which hinders the collapse of the current on breaking is the drawing out of an arc at the break. This can be prevented by placing a condenser *D* across the primary terminals which cuts off the flow through the arc, and therefore increases

$$\frac{di_1}{dt}.$$

Thus the electromotive force obtained from an induction coil is almost unidirectional, for the break of the circuit causes the main increase in secondary potential.

112. THE TRANSFORMER

The principle of a transformer is similar to that of the induction coil except that in this case the make and break of the induction coil is done away with, and the current fed into the transformer is a sinusoidal alternating current. The theory of the transformer is, however, very complicated owing to the effects of self-induction, mutual induction, and impedance on the oscillations. In these transformers, to increase efficiency, the magnetic circuit is generally complete, the windings being on a continuous ring of metal. Since the power put into a transformer in the primary is given by $E_1 i_1$, where E_1 is the potential across the primary and i_1 is the current in the primary, and since the output of power in the secondary can never exceed the input of power in the primary, we have the *limiting condition* that $E_2 i_2$, the electromotive force and current in the secondary, are related to those in the primary by the equation

$$E_2 i_2 = E_1 i_1$$

or

$$\frac{E_2}{E_1} = \frac{i_1}{i_2}.$$

If the efficiency of the transformer is Eff., then

$$E_2 i_2 = \text{Eff.}\ (E_1 i_1),$$

or

$$\frac{E_2}{E_1} = \text{Eff.}\,\frac{i_1}{i_2}.$$

The decrease in efficiency is due to a small loss caused by hysteresis effects in the iron core of the transformer, and while the loss is small the equation to be correct must contain the factor for this energy loss.

It can be shown from the theory that for open circuit work, or with a high resistance in the secondary, when there is a resistance in the primary, as is the case in induction coils, the ratio of E_2 to E_1 approaches the ratio n_2/n_1, where n_2 is the number of turns in the secondary and n_1 is the number of turns in the primary. This follows from the fact that the same flux passes through both primarv and secondary coils. Thus since the induced E.M.F. in the secondary is proportional to the flux and the number of turns, $E_2 = A\varphi n_2$, and the flux is proportional to the E.M.F., and number of turns in the primary (this holds only approximately as resistance and leakage reactance change this somewhat), $E_1 = A\varphi n_1$, therefore one can write

$$\frac{E_1}{E_2} = \frac{n_1}{n_2} \text{ or } E_2 = \frac{n_2 E_1}{n_1}.$$

For open circuit work the assumed conditions are met, and the law holds fairly well. As soon as a load is drawn the two equations assumed cease to hold with much precision and the statement is not strictly true. In general, the efficiency in the transformers in common usage is very high. The loss of energy in eddy currents and heating amounts to less than 4 per cent. Many transformers are 98 percent efficient.

CHAPTER XXIII

THE BALLISTIC GALVANOMETER—TRANSIENT PHENOMENA

113. THE BALLISTIC GALVANOMETER

An instrument of great importance in the measurement of electrical phenomena where currents change rapidly with time merits special consideration. This is the ballistic galvanometer. In Chapters XX and XXII, where induced currents were discussed, it was seen that the induced electromotive force generated was a function of the time rate of change of the current. In most cases this is not constant and is too rapid to follow by any instruments readily available (it can be studied by the oscillograph). The difficulty is overcome in many cases, as stated in Chapter XX, by letting the electromotive force generated cause a current to flow through the resistance coil of a galvanometer which is *little damped*, whose period is long compared to the duration of the current. Such a galvanometer is called a *ballistic* galvanometer. The instantaneous current i_{dt} flowing gives an instantaneous force f_{dt} which acts on the galvanometer coil for the time dt during which it flows. The product of f_{dt} times the time interval dt gives the impulse communicated to the coil at that instant. The sum of all the impulses over the time the current flows gives the total impulse given the coil, or the momentum communicated to it.

The sum of all the currents times the same time intervals dt gives the total quantity of electricity passing through the system resulting from the electromotive force generated. Thus one can write that $Q = \int_0^t i_{dt}dt$ and this is proportional to $\int_0^t f_{dt}dt$ or the momentum given the coil. The momentum causes the coil to swing to a point where the kinetic energy of rotation given to it by the field is changed into the potential energy of the twisted suspension. This deflection, or " throw," of the galvanometer is related in a definite way to the momentum given and therefore to the quantity passing.

Thus by measuring the " throw " of the galvanometer, one can measure Q and therefore determine some of the constants of the circuits studied in Chapters XX and XXII. In fact the ballistic gal-

vanometer is often used in determining the coefficient of self-induction and in a study of induced currents.

114. DERIVATION OF THE EQUATION OF THE BALLISTIC GALVANOMETER

In Chapter IX, the couple acting on the galvanometer coil of effective area A (area multiplied by the number of turns of wire), when a current i_{dt} flows through in a field H, was given by $G_{dt} = i_{dt}AH$. The force moment times the time dt over which the current has the value i_{dt} (the instantaneous impulse), is then given by

$$G_{dt}dt = i_{dt}HAdt.$$

Since the total impulse given the coil from the time 0 when the current starts to the time t when it has ceased is the sum of all the elementary impulses times their time of duration, one may write for the impulse

$$\overline{Gt} = \int_0^t G_{dt}dt = AH\int_0^t i_{dt}dt.$$

But $\int_0^t i_{dt}dt$ is nothing other than the total quantity of electricity which has passed, namely, Q: whence

$$\overline{Gt} = HAQ.$$

Now by Newton's second law, force times time is $ft = mat = mv$, mass times velocity, or momentum. By analogy one may write for the case where the force acts on a lever arm r, that $frt = mart = mvr$. But $v = \frac{ds}{dt}, \frac{ds}{dt} = \frac{rd\theta}{dt}$, and $vr = \frac{r^2d\theta}{dt}$, or $r^2\omega$, where ω is the angular velocity. This may be written $frt = mr^2\omega$. Now mr^2 for the simple case above, representing a mass m at a distance r, is the moment of inertia I. It accordingly follows that we can write that the impulse is $frt = I\omega$.

By analogy to this the couple times the time $\overline{Gt}$ is the impulse and must equal the moment of inertia of the coil times the angular velocity produced. Thus one may write

$$\overline{Gt} = I\omega = AHQ.$$

Again by Chapter IX, we have the galvanometer constant K' given by $K' = T_0/AH$, whence $AH = T_0/K'$,

$$I\omega = \frac{T_0Q}{K'}$$

and

$$Q = K'\omega \frac{I}{T_0}.$$

Now ω cannot be measured directly. But the quantity $\frac{1}{2}I\omega^2$ measures the kinetic energy of the coil just after the impulse has been given it. This causes the coil to deflect until the potential energy of the system equals the kinetic. The potential energy of the twisted suspension is the average torque times the angular displacement. At the rest point the torque is 0. At an angle θ it is $T_0\theta$ by Hooke's law. The average torque is then $\frac{T_0\theta + 0}{2} = \frac{1}{2}T_0\theta$. The displacement is θ. Thus the potential energy is $\frac{1}{2}T_0\theta^2$ and this must equal $\frac{1}{2}I\omega^2$.* Accordingly one has

$$\tfrac{1}{2}I\omega^2 = \tfrac{1}{2}T_0\theta^2$$

and

$$\omega = \theta\sqrt{\frac{T_0}{I}}.$$

Therefore

$$Q = K'\theta\sqrt{\frac{I}{T_0}}.$$

To get the ratio $\sqrt{\frac{I}{T_0}}$ one need only remember that the period of an oscillating system $T_1 = 2\pi\sqrt{\frac{I}{T_0}}$. (See Chapter IV.)

Thus the quantity Q of electricity which passed is given by

$$Q = \frac{K'\theta T_1}{2\pi}.$$

Thus from the " throw " θ of the ballistic galvanometer, its period of oscillation T_1 and its constant K', one can determine Q, the quantity which passed through its coil in a time short compared to its period. The quantity Q will be in coulombs if K' is measured in amperes per unit deflection. If K' is measured in absolute units, Q will be in absolute units.

* The equality stated applies to the maximum potential and maximum kinetic energies in oscillatory motion which occur at maximum and 0 deflection, respectively.

115. DAMPING

This equation is subject to one correction. The coil does not oscillate perfectly as a simple harmonic oscillator. That is, the second swing in the direction of the initial deflection is not quite as great, and each succeeding swing to one side is less than the preceding one. This is due to a frictional loss of energy in the suspension and to the action of the viscous drag of the air on the coil. The energy of oscillation is gradually converted into heat and the oscillation is said to be damped.

The result of this damping is that even the first throw of the galvanometer has not the true value which it would have had in the absence of damping. To calculate Q, the damping of the system must be known and the θ observed corrected for it.

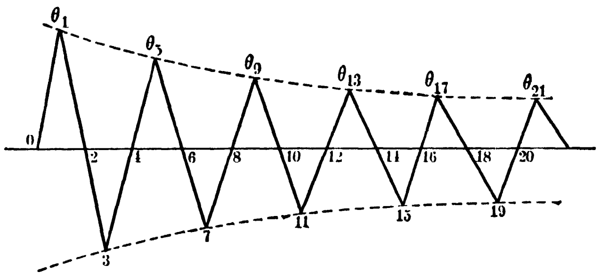


FIG. 112.—The Curve of a Damped Oscillation.

A simple harmonic oscillator has a displacement given by

$$\theta = \theta_0 \sin \frac{2\pi t}{T},$$

where T is its period, t the time at which the deflection is measured, θ_0 its maximum deflection, i.e., that when $t = T/4$ or an odd multiple of $T/4$. For a damped oscillation the equation becomes

$$\theta = \theta_0 e^{-pt} \sin \frac{2\pi t}{T}.$$

That is, the deflection θ at a time t is changed from the value it would have in an undamped oscillation by the quantity e^{-pt} where e is the base of the natural system of logarithms, t is the time and p is a constant. e^{-pt} is called the *damping factor*. It is seen that as t increases, e^{-pt} decreases indefinitely. That is, the throw at any time t will be diminished in the ratio of the ordinate of the exponential curve at t, plotted in Fig. 112, to unity; its value at $t = 0$. At successive quar-

ter swings, that is, from the rest position to its maximum deflection or the reverse, the times will be $t = 0$, $t_1 = \frac{T}{4}$, $t_2 = \frac{T}{2}$, $t_3 = \frac{3T}{4}$, $t_4 = T$, $t_5 = \frac{5T}{4}$, $t_6 = \frac{3T}{2}$, etc.; and the corresponding deflections θ_0, θ_1, θ_2, θ_3, θ_4, etc., will decrease as shown in the diagram of Fig. 112. The ends of the swings on either side will be on an exponential curve, $e^{-\rho t}$.

The ratio of any two successive maximum throws will then be given by

$$\frac{\theta_1}{\theta_3} = \frac{\theta_0 e^{-\frac{\rho T}{4}}}{\theta_0 e^{-\frac{3\rho T}{4}}} = e^{\frac{\rho T}{2}} = \frac{\theta_3}{\theta_5} = \frac{\theta_5}{\theta_7} = d.$$

The quantity $d = e^{\frac{\rho T}{2}}$ is called the *decrement* of the system, and $\log d = \log e^{\frac{\rho T}{2}} = \frac{\rho T}{2}$ is called the *logarithmic decrement.* Thus on any throw from 0 to an amplitude θ the value of θ observed must be multiplied by $\sqrt{d} = e^{\frac{\rho T}{4}}$. To get d, it is necessary to measure merely the value of the two successive throws on the same side of the rest point, that is say θ_1 and θ_5. Then

$$\frac{\theta_1}{\theta_5} = \frac{\theta_0 e^{\frac{-\rho T}{4}}}{\theta_0 e^{\frac{-5\rho T}{4}}} = e^{\rho T} = d^2.$$

Therefore

$$d = \sqrt{\frac{\theta_1}{\theta_5}}.$$

Since the difference in θ_1 and θ_5 for a good ballistic galvanometer is so small that the measurement is uncertain, it is common practice to take the amplitudes after say five successive complete swings. In the figure, with the notation used, this would mean measuring the amplitude of θ_1 and θ_{21}.

Then

$$\frac{\theta_1}{\theta_{21}} = \frac{\theta_0 e^{\frac{-\rho T}{4}}}{\theta_0 e^{\frac{-21\rho T}{4}}} = e^{\frac{20\rho T}{4}} = d^{10}.$$

Hence

$$d = \sqrt[10]{\frac{\theta_1}{\theta_{21}}}.$$

And to get true undamped θ_1, θ_1 observed must be multiplied by

$$\sqrt{d} = \sqrt[20]{\frac{\theta_1}{\theta_{21}}}.$$

This can be generalized so that for values of θ_n and θ_m where n and m are any whole numbers corresponding to the initial and final swings, properly counted, $d^2 = \sqrt[m-n]{\frac{\theta_n}{\theta_m}}$. The correct equation for the ballistic galvanometer is then given by

$$Q = \frac{K'T_1\theta}{2\pi}\left(\sqrt[m-n]{\frac{\theta_n}{\theta_m}}\right),$$

where θ is the first swing. It should be noted that Q will be in the system of units that K' has been measured for. If K' is in amperes, Q will be in coulombs.

116. TRANSIENT PHENOMENA

The definition of the derived unit, the coefficient of self-induction, enables us to investigate many new phenomena which have been inaccessible to us until now. To this class belong the phenomena known as *transient* electric phenomena. Since they involve the use of the concept of self-induction, they can help us to understand the nature of its action, as well as teach us new facts about electricity. It is therefore worth while to discuss two of these which are quite simple and illustrate admirably what happens in a circuit just after a contact is made or broken when self-induction or capacity are in the circuit and before the circuit reaches a steady state.

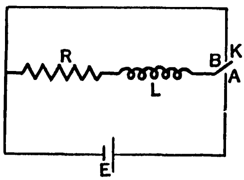


FIG. 113.—Circuit with Self-Induction and Resistance.

Case I. A Circuit with Self-induction and Resistance.—Regard the circuit pictured in Fig. 113 with a resistance R, a self-induction L, and an electromotive force E. Let the key be closed at A. The current begins to flow in the circuit. As it is building up, an E.M.F. of self-induction $-L\frac{di}{dt}$ is built up to oppose the growth of the current. The current was built up through the action of the impressed E.M.F., E. Thus we have $E - L\frac{di}{dt}$ as the E.M.F. in the circuit. By Kirch-

hoff's second law this at any instant must equal the iR drop in the resistance. Hence one can write

$$L\frac{di}{dt} + Ri = E.$$

This is what is known as a differential equation. To integrate it one may rearrange it as follows:

$$\frac{di}{\frac{E - Ri}{L}} = dt.$$

This may be transformed to

$$\frac{\frac{-L}{R}d\left(\frac{E - Ri}{L}\right)}{\frac{E - Ri}{L}} = dt, \text{ for } di = -\frac{L}{R}d\left(\frac{E - Ri}{L}\right)$$

The integral of this is

$$\frac{L}{R}\log\left(\frac{E - Ri}{L}\right) = -t + K,$$

in analogy with the operation

$$\int\frac{dx}{x} = \log x + k,$$

where K is a constant of integration. This K can be evaluated as follows. When $t = 0$, the key was open and $i = 0$. Putting these values into the equation,

$$\frac{L}{R}\log\left(\frac{E}{L}\right) = K.$$

Therefore

$$\frac{L}{R}\log\left(\frac{E - Ri}{L}\right) - \frac{L}{R}\log\frac{E}{L} = -t,$$

$$\log\frac{E - Ri}{E} = \frac{-Rt}{L},$$

or

$$\frac{E - Ri}{E} = e^{\frac{-Rt}{L}},$$

and

$$i = \frac{E}{R}\left(1 - e^{\frac{-Rt}{L}}\right).$$

This says that i, at any time t, is $i_0 = \frac{E}{R}$, the maximum value the current can have in the circuit when it reaches a steady state, multiplied by unity less a factor $e^{\frac{-Rt}{L}}$ which decreases with time and is never greater than unity. For small values of t, it is nearly unity and decreases with time to 0, so that the current rises from 0 at $t = 0$ to $\frac{E}{R}$ for the case that $t = \infty$. The curves obtained from this for various values of $\frac{R}{L}$ are seen in Fig. 114. The smaller $\frac{R}{L}$, the less the current rises for a given t. That is, the greater the self-induction relative to the resistance the more slowly the current rises. Thus L acts as an inertia tending to resist the establishment of the current. The inverse of $\frac{R}{L}$, that is, $\frac{L}{R}$ is called the *time constant* of the circuit. It determines how rapidly the current will change with t, and the time for a given change is greater, the greater the time constant.

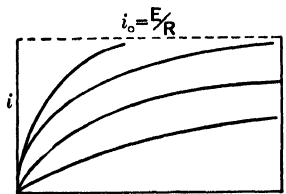


FIG. 114.—Rise of the Current in a Circuit Containing Self-Induction and Resistance.

If the key k had been closed at A and the current was flowing, breaking the key at A and connecting it to B would produce the following situation: at $t = 0$, i would be $\frac{E}{R}$. Then i would begin to die out, but an E.M.F. of self-induction $L\frac{di}{dt}$ would be generated trying to maintain the current. As opening the key removes the impressed E.M.F. one has $E = 0$. The current flowing i, and the E.M.F. of self-induction would both act in the same sense, as $L\frac{di}{dt}$ is trying to *maintain* the current. Thus

$$L\frac{di}{dt} + Ri = 0, \quad \text{and} \quad \frac{-L}{R}\frac{di}{i} = dt,$$

whence

$$\frac{L}{R}\log i = -t + K.$$

Now when

$$t = 0, i = \frac{E}{R} = i_0.$$

Thus

$$K = \frac{L}{R} \log i_0,$$

and hence

$$\log \frac{i}{i_0} = \frac{-R}{L} t, \quad \text{or} \quad i = i_0 e^{\frac{-R}{L}t} = \frac{E}{R} e^{\frac{-R}{L}t}.$$

It is seen that again the smaller $\frac{R}{L}$, the less i will decrease for a given t. So that again self-induction acts as an inertia tending to maintain the current. The curve of the decay of the current is seen in Fig. 115 for various values of $\frac{R}{L}$.

Case II. Charging a Condenser.—There is another simple case

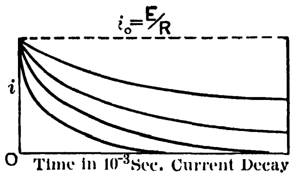


FIG. 115.—Decay of a Current in a Circuit Containing Self-Induction and Resistance.

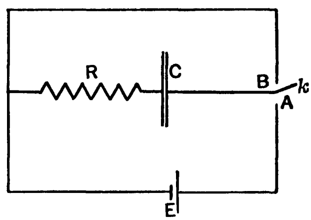


FIG. 116.—Circuit for Charging and Discharging a Condenser through a Resistance.

which merits treatment at this point. It does not involve the use of self-induction, but indicates in an interesting fashion the action of a condenser on transient currents. Consider the circuit pictured in Fig. 116. The condenser C is in series with a resistance R and a battery E. k is a key which, to begin with, can be considered open. If it makes contact at A, the battery E charges the condenser. If it makes contact at B, the condenser discharges through R. The charge on the condenser is 0 at the instant the key k is closed at A. Then a current begins to flow charging the condenser up. In charging up the condenser, an E.M.F. opposing the flow of the current is built up. The back E.M.F. due to the charge on a condenser is given by the

relation $E' = \frac{Q}{C}$, where Q is the quantity and C is the capacity. Thus Kirchhoff's second law may be expressed as

$$E - E' = E - \frac{Q}{C} = Ri.$$

Now

$$Q = \int i dt, \quad \text{or} \quad i = \frac{dQ}{dt}.$$

Therefore

$$R\frac{dQ}{dt} + \frac{Q}{C} = E.$$

Algebraic transformation permits us to write

$$\frac{-CRd\left(\frac{E - Q/C}{R}\right)}{\frac{E - Q/C}{R}} = dt$$

and integration yields

$$CR \log \left(\frac{E - Q/C}{R}\right) = -t + K.$$

As $Q = 0$ at $t = 0$,

$$CR \log \frac{E}{R} = K$$

and

$$\log \frac{E - Q/C}{E} = -t/CR$$

or

$$Q = EC\left(1 - e^{\frac{-t}{CR}}\right).$$

But $EC = Q_0$, the final charge on the condenser; and we can write

$$Q = Q_0\left(1 - e^{\frac{-t}{CR}}\right).$$

This says that the charge on the condenser at any time t is Q_0, the final charge when $t = \infty$, less $Q_0e^{\frac{-t}{CR}}$. That is, the greater C and the greater R, the more slowly does the condenser charge up.

For the case where the condenser is charged and the battery is removed, letting the condenser discharge through a resistance R, one

throws k from A to B. In this case E is 0 and one had the E.M.F. $E' = \frac{Q}{C}$. Then as the only other E.M.F. in the circuit is the iR drop, the equation is the same as before with $E = 0$. Hence

$$Ri + \frac{Q}{C} = 0, \quad \text{or} \quad R\frac{dQ}{dt} + \frac{Q}{C} = 0,$$

whence

$$\log Q = -\frac{t}{CR} + K,$$

and as $Q = Q_0$ at $t = 0$, $K = \log Q_0$. Thus

$$\log\frac{Q}{Q_0} = \frac{-t}{CR} \quad \text{and} \quad Q = Q_0 e^{\frac{-t}{CR}}.$$

The charge therefore falls off exponentially with the time and falls the more slowly the larger C and R. Thus the condenser acts to *delay the arrival of a steady state where quantity is involved.* It causes *potential* to *build up slowly* and to disappear slowly.

It is of interest to see how the *current* varies in this case. The current equations may be derived at once from the quantity equations above. Since $i = \frac{dQ}{dt}$, one has for the case of charging a condenser

$$i = Q_0\frac{d}{dt}\left(1 - e^{\frac{-t}{CR}}\right) = \frac{Q_0}{CR}e^{\frac{-t}{CR}},$$

and for discharge,

$$i = Q_0\frac{d}{dt}\left(e^{\frac{-t}{CR}}\right) = -\frac{Q_0}{CR}e^{\frac{-t}{CR}}.$$

As $E = \frac{Q_0}{C}$, then

$$\frac{Q_0}{CR} = \frac{E}{R} = i_0.$$

Hence for charge one has

$$i = i_0 e^{\frac{-t}{CR}},$$

and for discharge,

$$i = -i_0 e^{\frac{-t}{CR}}.$$

These equations are plotted beside the equations for quantity (see Figs. 117*A* and *B*). It is seen that while the *quantity* on charge rises from *zero* to a *saturation* value, the *current* causing the charging *falls*

off from an *initial value* i_0 to zero after a long time. That is, the current starts at a maximum, while the charge on the condenser is zero, and falls to zero as the condenser charges up.

In the case of discharge the current i_0 *has a negative sign*, signifying that it flows in the opposite sense to the positive charging current. As before it is a maximum at the start and goes to zero just as the quantity Q falls to zero. It is important to note this difference in behavior of the charge and current. It is an illustration of a fact to be observed in Chapter XXIV, that *capacity advances the phase of a current, while self-induction retards it,* for i is greatest at $t = 0$ in the case of capacity, while i is least when $t = 0$ when self-induction is present. The charging current with capacity is therefore a maximum

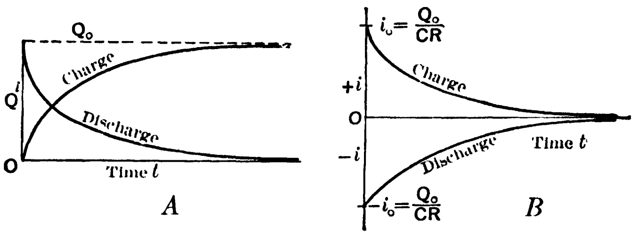


FIG. 117.—Curves of Charge and Discharge of a Condenser Through a Resistance; *A*. Variation of *Quantity* on the Condenser with Time, *B*. Variation of the Charging or Discharging *Current* with Time.

at the beginning and is ahead of the charging current, which is zero for the case when induction is present, when the switch is closed. Thus capacity is said to cause a lead of the current while inductance is said to cause a lag in the growth of the current.

117. MEASUREMENT OF HIGH RESISTANCES

In the case of resistance and self-induction where the equation for the dying out of a current is $i = i_0 e^{-\frac{R}{L}t}$, it is seen that if $\frac{R}{L}$ is great enough one might expect to observe the change in i with time. By using a specially designed galvanometer of very short period (vibration galvanometer) one can actually follow the decay of the current.

In the case of the discharge of the condenser, the equation found was

$$\frac{Q}{Q_0} = e^{\frac{-t}{CR}}.$$

If C and R are very large, $\frac{Q}{Q_0}$ will vary with time in a measurable

amount. For a C of 1 microfarad and $R = 10^8$ ohms (Toulene resistance), the time of discharge would lie in the neighborhood of a minute or two. As C can be easily measured, and as resistances of 10^8 ohms cannot be easily measured, this decay is used as a means of measurement of high resistances. The circuit is pictured in Fig. 118. E is a battery for charging the condenser C of known capacity which has little dielectric absorption. R is the unknown resistance and G is a galvanometer of the ballistic type. k_2 is a charging key and k_1 is the key used in measurement. To charge, k_1 is open and separated from both A and B, while k_2 is closed. k_2 is then opened and k_1 thrown to B. The condenser discharges through G and the ballistic throw measures Q_0, the charge on the condenser. Then the condenser is charged again. This time the key k_1 is closed to A for a suitable time t, say 30 seconds. Then it is rapidly switched to B and the deflection of G noted. This gives Q the quantity left on the condenser after a time t. Having Q, Q_0, t and C, one can at once solve for R.

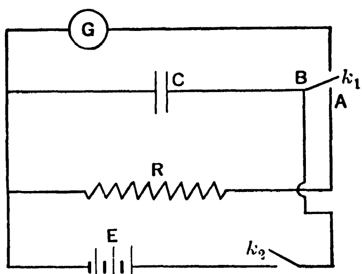


FIG. 118.—Measurement of High Resistance by the Discharge of a Condenser through it.

It might be noted that the factor $e^{-\frac{R}{L}t}$ and $e^{\frac{-t}{RC}}$ must be dimensionless as $\frac{i}{i_0}$ is a mere number. Thus $\frac{R}{L}$ and $\frac{1}{RC}$ must have dimensions of $\frac{1}{T}$. For this reason it is important to remember that R and L, and R and C must be expressed in the same system of units, and such as are consistent with the dimensions above.

118. CATHODE RAY OR BRAUN TUBE OSCILLOGRAPH

The study of transient phenomena and the analysis of electrical currents has been carried out in recent years by means of the cathode ray oscillograph. The use of the oscillograph has become so universal that it seems worth while to digress a moment to consider the mechanism of this instrument in connection with the study of transient phenomena. In Chapter XXVI, it will be discovered that incandescent metals emit a stream of electrons if the temperature is high enough. In a good vacuum such a stream of electrons can be accelerated in an electrical field between the hot filament F (see Fig. 119)

to a small pinhole in a diaphragm D by placing a potential of from 300 to 5000 volts between the diaphragm D and the filament F, making the diaphragm positive. The electrons which strike the opening have acquired such a high velocity that they are able to pass through the opening without striking the diaphragm. They then traverse the rest of the tube and strike on a photographic plate P. The oscillograph in the space between D and P is also provided with two pairs of small plate electrodes, E and E', sealed into the space between D and P. One pair of the plates, E', have their planes vertical and their long axes parallel to the cathode ray beam. The other pair, E, have their planes horizontal and their long axes parallel to the cathode ray beam. The tube is also so arranged that it can be placed inside of a uniform magnetic field H, perpendicular to the plane of the paper and the pole, pieces of which are indicated by the circle I. This equip-

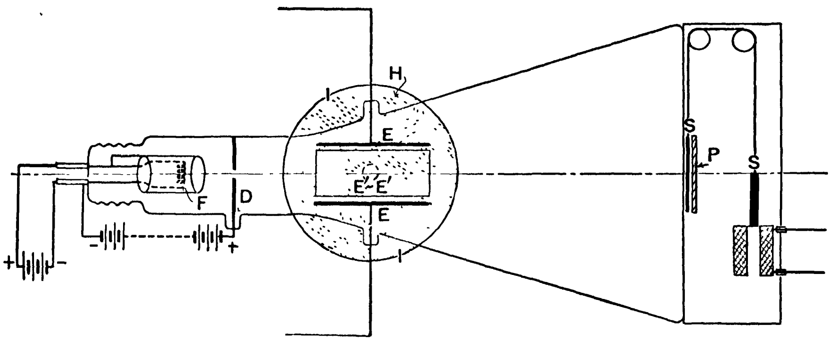


FIG. 119.—Cathode Ray, or Braun Tube Oscillograph.

ment will allow a large number of different studies to be made with the instrument. In the absence of any fields the beam of cathode rays emerging from the pinhole travel in a straight line and impinge at a point on the photographic plate P which can be exposed by opening a shutter S in front of it. This stream of electrons constitutes a negative current of electricity flowing from D to P or a current of positive electricity flowing from P to D. If the cathode ray beam is subjected to an electric field between the vertical plates E' which increases in a uniform fashion linearly with time a straight line is traced along the photographic plate from the undeflected position indicating the zero position of the cathode ray beam, into or out of the plane of the figure. If the rate of increase of the field between E' is known we can for equal lengths of the film represent equal time intervals along this line. If it is desired to measure transient currents the current to be studied is run through the windings of the coil acti-

vating H producing a field at right angles to the beam of electrons. This beam acts to deflect the cathode ray beam up or down, depending on the direction of the current through the coil. The deflection will be proportional to the field produced by the current and thus proportional to the current. Consequently as the current varies with time the vertical deflection of the cathode ray beam from its zero position will vary and the variation is automatically drawn out on the time axis by setting the increasing field across the plates E' into operation. By an elaborate electro-mechanical system the field between plates E' and the field accelerating the electrons can be applied by cutting out the proper portion of a single surge of a sine wave of known amplitude and frequency. By properly centering the spot produced by the cathode ray beam on the photographic film and making marks representing the time intervals on the film, the film on development will yield accurate curves describing what happened in the discharge through the coil. The changes of potential can be studied equally conveniently by placing the transient source of potential across the two horizontal plates E. Any changes in potential between these plates will cause a change in the deflection of the beam which, drawn out on a time axis by the field between plates E', again plots the curve of the occurrence. In the study of simple harmonic motions it is sometimes not necessary to proceed in this manner with a photographic film inasmuch as the phase and relative amplitudes of two simple harmonic motions at right angles can be determined from the dimensions of the curves traced by the superposition of these two motions. It is simply necessary to superimpose the two potentials to be studied on the two pairs of plates at right angles to the cathode ray beam and to each other. Because of the persistance of vision and phosphorescence the Lissajou figure described on a phosphorescent zinc sulphide screen replacing the photographic film at the end of the tube can be observed and studied at will.

Owing to the very small mass of the electron, by properly regulating the velocity of the beam the response of the beam can be made to suit almost any conditions and the time required to deflect the beam is as short as almost any time physically attainable. The chief limitation in time measurement occurs through the limitation of the speed of the uniformly increasing P.D. generating the time axis. Owing to the advance in the technique of high frequency oscillations the time axis can be drawn out so as to represent exceedingly short intervals of time. In recent years, Rogowski has actually succeeded in studying the transient phenomena in the spark over time intervals of the order of 10^{-7} seconds.

CHAPTER XXIV

ALTERNATING CURRENTS I. DEFINITIONS, ROOT MEAN SQUARE AVERAGES, AND EFFECT OF SELF-INDUCTION

119. THE IMPORTANCE OF ALTERNATING CURRENTS

In the last chapter it was shown that self-induction became of importance when currents vary with time, that is, self-induction causes an electromotive force only when we have a time rate of change of current,

$$\frac{di}{dt}.$$

Besides the transient effect of the last chapter, which emphasizes this, there is another important group of currents, those altering periodically with time.

Now it seems strange that one should have to deal with alternating currents when direct currents can be generated quite simply. However, in modern industrial life efficiency demands that electrical power be generated in central plants. In these plants the distribution of power to outlying points is a great problem. The loss in power distribution comes almost entirely in the i^2R loss in the line. Where electrical power is transmitted over 20 miles of line and as it is today over hundreds of miles, R is quite a factor and the currents range in the thousands of amperes. Thus the i^2R loss in heating is very great, and for transmission of direct currents the loss would be dependent on this factor. However, in alternating currents it is possible to transmit the alternating current at a very high voltage, that is, E of the term Ei which expresses the power sent is large, and i correspondingly small. By means of transformers which we have discussed in the last chapter, it is possible with very little loss of energy to convert the high potential alternating current to a low potential at a substation. Thus today practically all power transmission depends on the use of alternating currents. Again, in all telephony and especially in radio, we are dealing with changes produced in the electrical system by

periodic changes in currents. It is consequently essential to study the alternating current.

120. DEFINITION OF ALTERNATING CURRENT TERMS

In Chapter XXI it was shown that the simple dynamo gives an E.M.F. $E_0 = \frac{2\pi HAN}{10^8} \sin \theta$. One may set $\frac{2\pi HAN}{10^8} = E_0$, a constant of the generating end of the circuit, and the equation becomes $E = E_0 \sin \theta$. The angle θ represents the angle at which the generator coil is at the instant when it is generating the particular E.M.F. considered. As this angle varies with the time, it is seen that the E.M.F. varies with the time. If we plot E as a function of the angle, we will see that we have a curve of the type indicated in Fig. 120. This curve is a sine curve and it is seen that when θ takes on the values

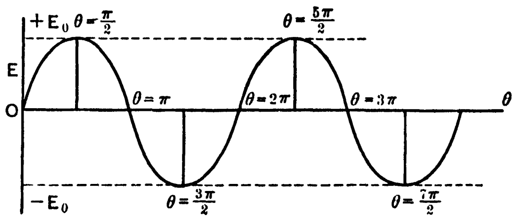


FIG. 120—Plot of a Sinusoidal E.M.F.

$$\frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}$$

or any odd multiples of $\frac{\pi}{2}$ radians, the curve has either a maximum or a minimum. We know that the sine of $\frac{\pi}{2}$ radians or 90° is equal to unity, therefore at these points

$$E = +E_0 \quad \text{or} \quad -E_0.$$

It is thus seen that E_0 represents the *peak values of the E.M.F.* generated, and it is often termed the *amplitude* of the oscillation. Now the angle θ gives the angle of the armature with regard to the field of the dynamo. If we had started a dynamo from a position in which the armature was not at the zero point as in Fig. 101*C*, but had started with an initial angle θ which we now will designate as φ, the curve of Fig. 120 would have been shifted over to the right or the left of the zero so that when $\theta = 0$ in the figure the position of the curve on the E axis corresponded to the angle φ of the armature. This angle φ is called *the phase* of the alternating potential.

It is more useful to regard the alternating potential not from the

point of view of the *angle* of the armature, but from the point of view of *the time which has elapsed after the armature was in its initial position represented by the angle φ.* It is obvious that once in every revolution of the armature, the armature will be back at its initial phase φ. That is, the quantity representing the position of the armature relative to φ on the time scale must be of such a nature that it repeats itself over the revolution. Now a revolution in angular measure equals 2π radians. Thus if we designate by t the time which has elapsed since the initial angle φ, which marks the beginning of our time reckoning, we would represent the periodical variation by the term

$$\frac{2\pi t}{T},$$

where T represents the period of oscillation, that is, the time for the armature to make one complete revolution. It will be seen at once that the expression

$$\frac{2\pi t}{T}$$

$\left(\text{the fractional part } \frac{t}{T} \text{ of a whole revolution}\right)$, can correctly represent the angle in radians through which the armature has turned in a time t. For every time that t is equal to T, the angle becomes 2π radians. Thus the quantity which we called θ representing the angle of the armature may now be represented by the expression

$$\theta = \frac{2\pi t}{T} + \varphi.$$

We may therefore write the expression for the E.M.F. as

$$E = E_0 \sin\left(\frac{2\pi t}{T} + \varphi\right).$$

The quantity T representing the time of a single revolution is the *period* of the oscillation. Sometimes it is simpler to count the number of revolutions in a second, N. Since T is equal to $\frac{1}{N}$, the expression above may be written

$$E = E_0 \sin(2\pi N t + \varphi).$$

We will recapitulate the names of the quantities concerned in these two equations.

In the above equations:

E is the *instantaneous value of the E.M.F.*
E_0 is the *amplitude of the oscillation* (maximum value).
φ is the *phase of the E.M.F.*
T is the *period.*
N is the *frequency of oscillation.*

121. VALUES OF CURRENTS GIVEN BY INSTRUMENTS, ROOT MEAN SQUARE CURRENT

The next question which arises in these considerations is the following: Suppose we have an E.M.F. of the type discussed above, having, as is the case in lighting circuits, the frequency of 60 cycles per second, what will our measuring instruments record with such a current? In the first place if we are going to measure alternating currents we cannot measure such currents or potentials with ammeters and voltmeters having fixed magnetic fields as indicated in Chapter IX, for every time that the current or potential reverses the force acting on the moving coil of the instrument will reverse, and if the instrument could not follow the rapid alternations of the current or potential, no reading would be obtained. If, however, we let the current which flows through the moving coil also excite the magnetic field, the direction of the torque on the moving coil will always be the same and a *varying deflection* proportional to current or potential squared will be obtained provided that the coil has so small an inertia that it can follow the rapid oscillations. Actually this is not the case and the coil takes up a position which is intermediate between that of the zero position and the maximum position. That is, *all alternating current potential and current measuring devices give a deflection which is neither the peak, or maximum, amplitude of the quantity nor the zero value.* In fact, we deal with an average E.M.F. or current. This must be related in some fashion to the maximum value, and the relationship between the average and the maximum value of E must be known in order to apply our equation for the E.M.F. at any instant.

In deducing this value, we must compute the average E.M.F. or current. To do this, we must remember, however, that since the same current flowing through the coil and through the field is twice acting to cause the deflection, *the deflection is proportional to the product of the current by itself, or to i^2.* The same applies to the E.M.F. The problem before us thus is to calculate the average square of the current for a current which varies sinusoidally with the time. The deduction to be given for the current is the same as would be given for the

E.M.F., for, barring the effect of self-induction to be studied later, we have

$$i = \frac{E}{R}$$

by Ohm's law. The average square of the current $\bar{i^2}$ is obtained as follows. From the law of the alternating potential, we can write that

$$\frac{E}{R} = \frac{E_0 \sin \theta}{R}$$

and thus

$$i = i_{max} \sin \theta.$$

Now we wish to obtain the average value of the square of this current. This is equivalent to the following graphical problem. In Fig. 121, we wish to find the height of the rectangle of base π whose area is equal to the area under the $\sin^2$ curve, for we are dealing with the squares of the current. The height of this rectangle would be the height of the equivalent steady current $\bar{i^2}$, which would have the same area, or the same effect on the needle, as the alternating current has. We are interested only in the area of a single positive loop of the current for the simple reason that the negative loop due to the reversal of the current in the field coils produces a magnetic force in the same direction as the initial force. Now the area of such a rectangle is $\bar{i^2}\int_0^\pi d\theta$, or $\bar{i^2}$ times π, its base. The area under one loop of the $\sin^2$ curve would be obtained as follows. The height of any element $d\theta$ wide along the base is $i^2_{max} \sin^2\theta$. Its area is $i^2_{max} \sin^2\theta d\theta$. The sum of all these elements $d\theta$ from 0 to π gives the total area. Since by definition the area of the rectangle must equal this area, one can write

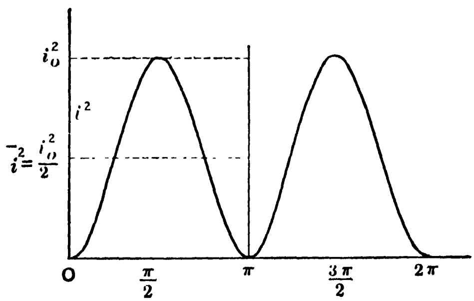


FIG. 121.—Determination of the Average Squared Current for a Sinusoidal Alternating Current.

$$\bar{i^2}\int_0^\pi d\theta = \bar{i^2}\pi = \int_0^\pi i^2_{max} \sin^2 \theta d\theta.$$

Integration of this leads at once to the value

$$\overline{i^2} = \frac{i^2_{\max}}{2}.$$

Hence the

$$\sqrt{\overline{i^2}} = \frac{i_{\max}}{\sqrt{2}},$$

and in the same fashion since R is a constant not involved in integration,

$$R\sqrt{\overline{i^2}} = \frac{Ri_{\max}}{\sqrt{2}}$$

or

$$\sqrt{\overline{E^2}} = \frac{E_0}{\sqrt{2}}.$$

We therefore see that the average current or the average E.M.F. which is registered by an alternating current instrument is the maximum current or potential divided by the $\sqrt{2}$. These average currents and potentials registered by instruments are called the *root mean square* values of the current and potential; and when speaking of the currents and potentials in amperes and volts, they are designated in engineering practice by *virtual amperes and virtual volts.*

122. EFFECT OF SELF-INDUCTION ON AN ALTERNATING E.M.F.

The next question which we wish to consider is the effect of the self-induction in a coil on an alternating current. It is obvious that if we have an alternating current in a coil whose self-induction is L,

$$L\frac{di}{dt}$$

has a finite value when $\frac{di}{dt}$ has a finite value. Assume that $i = i_0 \sin\frac{2\pi t}{T}$, and hereafter represent $\frac{2\pi}{T}$ by p. Then $i = i_0 \sin pt$, and

$$\frac{di}{dt} = + i_0 p \cos pt.$$

Thus there will be induced in the coil an E.M.F. E' of self-induction whose value

$$E' = L\frac{di}{dt} = Li_0 p \cos pt = Li_0 p \sin\left(pt + \frac{\pi}{2}\right).$$

Thus, the E.M.F. of self-induction will be $+ 90°$ or $+ \pi/2$ radians out of phase with the current which causes it. The resultant E.M.F. in the circuit of Fig. 122 can be computed at once from Kirchhoff's second law. The E.M.F. existing in the circuit at any time t is the sum of the iR drop plus the E.M.F. of self-induction. Thus we may write for the circuit of Fig. 122, $E - E' = Ri$. Since $i = i_0 \sin pt$ and $E' = Li_0p \cos pt$ the equation for the *instantaneous* value of the E.M.F. becomes $E = Ri_0 \sin pt + Li_0p \cos pt$.

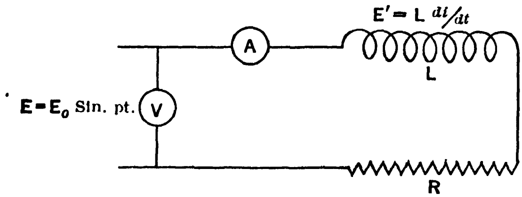


FIG. 122.—Circuit for the Study of the Effect of Self-Induction on an Alternating E.M.F.

As a rule we are interested in the effective or average value of the E.M.F. as this is what is read by the instruments used. We will therefore find that it is important to calculate the value of this average in order better to understand the foregoing equation. As in the case of the currents previously discussed, the average squared voltage over an interval of time dt taken for half a cycle (that is, from 0 to $\frac{T}{2} = \frac{\pi}{p}$) must equal the area under the curve represented by the value of E^2 which was derived for a circuit having self-induction. We must therefore set the integral of $E = Ri_0 \sin pt + Li_0p \cos pt$ squared and multiplied by dt equal to the average E_v^2 times the integral of dt from 0 to $\frac{\pi}{p}$.

Thus we write

$$E^2 = R^2i_0^2 \sin^2 pt + L^2i_0^2p^2 \cos^2 pt + 2Li_0^2pR \sin pt \cos pt$$

and

$$E_v^2\int_0^{\frac{\pi}{p}} dt = \frac{\pi}{p}E_v^2 = R^2i_0^2\int_0^{\frac{\pi}{p}} \sin^2 ptdt + L^2i_0^2p^2\int_0^{\frac{\pi}{p}} \cos^2 ptdt$$
$$+ 2Li_0^2Rp\int_0^{\frac{\pi}{p}} \sin pt \cos ptdt.$$

Integration makes the last term 0 for each limit and the first two terms are $\frac{\pi}{2p}$. Thus we have as a result the *important* equation.

$$E_v^2 = R^2\frac{i_0^2}{2} + L^2p^2\frac{i_0^2}{2}.$$

This is the same as

$$E_v^2 = R^2 i^2_v + L^2 p^2 i_v^2$$

and also leads to the result

$$E_0^2 = R^2 i_0^2 + L^2 p^2 i_0^2.$$

The significance of this relation is seen at once by solving the equation for

$$i_v = \frac{E_v}{\sqrt{R^2 + L^2 p^2}}.$$

This says that the current i_v measured by an A.C. ammeter A in Fig. 122 in the circuit is not the current found by dividing E_v by R, as the simple Ohm's law would lead one to believe, but it contains in addition a term Lp which is called the *inductive reactance* of the circuit, depending on L and p. It is seen that the greater L or the greater p (i.e., the frequency) the more the law departs from Ohm's law, the current being less than it would be on that law. The quantity $\sqrt{R^2 + L^2 p^2}$ is called the *impedance and is designated by the letter z.* Thus the more general form of Ohm's law is

$$i = \frac{E}{z} = \frac{E}{\sqrt{R^2 + p^2 L^2}}$$

which reduces to Ohm's law for $p = 0$, that is, for $N = 0$ or $T = \infty$.

This behavior can be readily shown by the circuit of Fig. 122 having a value of L of the order of 1 or 2 henrys and an R of some 100 ohms using an electric light globe in place of the ammeter A. When a D.C. source of potential is applied the lamp will glow brightly, while with an A.C., it will hardly glow at all for a frequency of 60 cycles.

The equation for the instantaneous value of the electromotive force E,

$$E = Ri_0 \sin pt + Li_0 p \cos pt$$

has an awakward form and by the procedure to be outlined can be thrown into a far more useful and significant form. Let us designate for simplicity Ri_0 by X, and $Li_0 p$ by Y. We can then lay off the values of X and Y along the axes x and y of a rectangular coordinate system, shown in Fig. 123, and the equation above takes on the form

$$E = X \sin pt + Y \cos pt.$$

Draw through the end of Y a line parallel to X and through the end of X a line parallel to Y and join 0 to the corner of the rectangle so formed. The diagonal C of the rectangle makes an angle φ with X and has the following properties:

$$Y = C \sin \varphi$$

$$X = C \cos \varphi$$

$$X^2 + Y^2 = C^2.$$

The expression for E then takes the form $E = C \cos \varphi \sin pt + C \sin \varphi \sin pt$ which equals $C \sin (pt + \varphi)$ by a well known trigonometrical relation. We now have E in a much more convenient form in that it is a *single* trigonometric function of pt containing two constants, C and φ. $E = C \sin (pt + \varphi)$. To identify these constants, we proceed as follows: $C^2 = X^2 + Y^2 = R^2 i_0{}^2 + L^2 i_0{}^2 p^2 = E_0{}^2$, whence $C = E_0$, the maximum amplitude of the *impressed* E.M.F., and E is given by

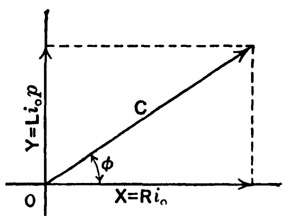


FIG. 123.—Diagram for Transformation of the Equation for the Instantaneous Value of an A.C. in a Circuit with Self-Induction and Resistance.

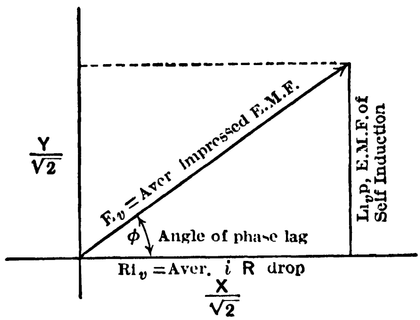


FIG. 124.—Vector Diagram for an A.C. Circuit with Self-Induction and Resistance.

$E = E_0 \sin (pt + \varphi)$. Again we may remember that $i = i_0 \sin pt$ and we see that the *phase of the E.M.F., E, is φ radians or degrees, ahead of the current. The current lags behind the E.M.F. in phase by an angle φ.* Its value can be found at once from the relation

$$\tan \varphi = \frac{Y}{X} = \frac{Li_0 p}{Ri_0} = \frac{Lp}{R}.$$

Thus besides having the amplitude reduced below what one would expect for a self-induction equal to 0 or for a steady current (i.e., the use of impedance in place of ohmic resistance) the current lags behind

the E.M.F. an amount depending on the size of $\frac{Lp}{R}$. It is seen that as in transient currents there is a factor depending on the ratio L/R which marks the lagging of the current due to self-induction. For large values of $\frac{Lp}{R}$, $\tan \varphi$ approaches infinity, and φ approaches 90° or $\pi/2$ radians. This means that the E.M.F. is a maximum when i is 0 and vice versa, and leads to interesting results to be discussed in the next chapter. A similar consideration will be applicable in the case of capacity. The results of this discussion can be pictured by means of the following diagram. Fig. 124 illustrates the usual vector relation used by engineers and follows directly from the representation of X and Y on a rectangular coordinate axis.

CHAPTER XXV

ALTERNATING CURRENTS II—EFFECT OF CAPACITY ON ALTERNATING CURRENT AND THE EFFECT OF CAPACITY AND SELF-INDUCTION COMBINED—ELECTRICAL OSCILLATIONS

123. EFFECT OF CAPACITY ON THE IMPEDANCE OF AN ALTERNATING CURRENT

In the last chapter the effect of self-induction on an alternating current was studied. When we attempt, however, to carry the same mode of approach to the action of capacity on a circuit having an impressed alternating sine wave form E.M.F. we encounter difficulties which make the mode of approach used for self-induction impossible. Since E', the back E.M.F. produced by capacity, is given by $E' = Q/C$, where Q is quantity and C is capacity, and since $i = \frac{dQ}{dt}$, the expression for E' becomes $E' = \frac{\int idt}{C}$. In choosing the limits of integration in order to compare the equation with the current equation, we find a troublesome constant of integration entering in, justification for the dropping of which cannot be made in this form. We must therefore proceed in a different manner, and this is one quite parallel to that used in studying the effect of capacity on transient phenomena.

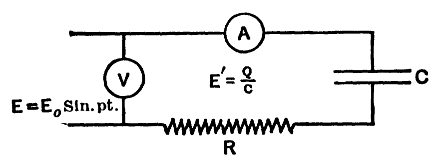


FIG. 125.—Circuit for the Study of the Effect of Capacity and Resistance on a Sinusoidal A.C.

Let us assume a circuit of the form shown in Fig. 125 and that the impressed E.M.F. has the amplitude $E = E_0 \sin pt$. Then Kirchhoff's second law leads to the expression $E - E' = Ri$, where E' is the back E.M.F. due to C, and R is the resistance in the circuit, while i is the current. Now $E' = Q/C$, where Q is the quantity of electricity on C at any time and $i = \frac{dQ}{dt}$. Hence we can write

$$R\frac{dQ}{dt} + \frac{Q}{C} = E_0 \sin pt,$$

whence

$$\frac{dQ}{dt} + \frac{1}{RC}Q = \frac{E_0}{R}\sin pt.$$

This equation is a linear differential equation of the first order with constant coefficients. The general solution of this is found in any text on differential equations and is as follows:

$$Q = \frac{E_0}{R}\,\frac{\left(\frac{1}{RC}\sin pt - p\cos pt\right)}{\left(\frac{1}{RC}\right)^2 + p^2} + ce^{-t/RC}.$$

The first member on the right-hand side of the equation is a term that varies periodically with t and is the one of importance in this problem. The second term multiplied by the undetermined constant c represents a term that approaches 0 as t becomes large. It is the term that represents the *transient* effect when the switch is first thrown and is in fact quite similar to the transient term that we found in Chapter XXIII for capacity. As we are interested in the situation after this has died out (i.e., we measure currents some seconds after the switch has been closed), we can set t very great and this term approaches 0. We must now throw the equation into a more convenient form for study. To do so let us write, as in Chapter XXIV, that $1/RC = B \sin\varphi$ and $p = B\cos\varphi$, where $B^2 = \frac{1}{R^2C^2} + p^2$.

Then

$$Q = \frac{E_0}{RB}\,\frac{(\sin\varphi\sin pt - \cos\varphi\cos pt)}{\sin^2\varphi + \cos^2\varphi},$$

whence

$$Q = \frac{-E_0}{RB}\cos(pt + \varphi).$$

Now

$$i = \frac{dQ}{dt} = \frac{pE_0}{RB}\sin(pt + \varphi),$$

or

$$i = \frac{E_0}{\frac{R}{p}\sqrt{\frac{1}{R^2C^2} + p^2}}\sin(pt + \varphi),$$

therefore

$$i = \frac{E_0}{\sqrt{R^2 + \frac{1}{p^2C^2}}} \sin (pt + \varphi),$$

which is the equation sought for the current in this circuit. From this we can also write, since $i = i_{max}$ when $\sin (pt + \varphi) = 1$,

$$i_{max} = \frac{E_0}{\sqrt{R^2 + \frac{1}{p^2C^2}}}, \text{ and } i_v = \frac{E_v}{\sqrt{R^2 + \frac{1}{p^2C^2}}}.$$

Since $\frac{1}{RC} = B \sin \varphi$ and $p = B \cos \varphi$,

$$\frac{\sin \varphi}{\cos \varphi} = \tan \varphi = \frac{1}{RpC}.$$

Since we chose $E = E_0 \sin pt$, we see that *the current i is ahead of E by a phase angle φ.* In the case of self-induction the phase angle φ was the angle by which the *E.M.F. was ahead of the current.* Thus we concluded that self-induction causes a *lag in the current behind the impressed E.M.F.*, while with capacity the *current leads or is advanced in phase ahead of the impressed E.M.F.* Again as before the angle of phase advance φ can be found at once from the relation $\tan \varphi = \frac{\sin \varphi}{\cos \varphi} = \frac{1}{RpC}$. We also note that in this circuit the virtual current does not follow Ohm's law, but that

$$i_v = \frac{E_v}{\sqrt{R^2 + \frac{1}{p^2C^2}}}.$$

That is, the effect of capacity is to reduce the current below that to be expected when it is absent. The capacity impedance is $\sqrt{R^2 + \frac{1}{p^2C^2}}$ and the capacity reactance $1/pC$. The greater C and p, the more nearly is Ohm's law obeyed, while the smaller C and p, the greater is the departure. This can be readily shown by placing either a small capacity or a large one in an alternating circuit with an electric light in series. While 10 microfarads at 60 cycles lights the lamp, 1 microfarad does not do this. This use of $\sin \varphi$ and $\cos \varphi$ for $1/RC$ and p of the equation leads one to be able to make the same sort of a vector diagram for the effect of capacity as for self-induction.

124. CASE WHERE SELF-INDUCTION, CAPACITY AND RESISTANCE ARE PRESENT

In the case where both of these quantities are in the circuit, the differential equation becomes

$$E - (E'_1 + E'_2) = Ri,$$

where E'_1 is the E.M.F. of self-induction and E'_2 is the E.M.F. due to capacity. On setting $E = E_0 \sin pt$, we obtain

$$L\frac{di}{dt} + Ri + \frac{Q}{C} = E_0 \sin pt,$$

and

$$\frac{d^2Q}{dt^2} + \frac{R}{L}\frac{dQ}{dt} + \frac{Q}{LC} = \frac{E_0}{L}\sin pt.$$

The resulting equation is more difficult to solve, but it leads to a solution of the same form as for capacity alone with

$$i = \frac{E_0}{\sqrt{R^2 + \left(\dfrac{1}{pC} - Lp\right)^2}} \sin (pt + \varphi)$$

where φ is now defined by the equation

$$\tan \varphi = \frac{\dfrac{1}{Cp} - Lp}{R}.$$

This equation is quite analogous to the one for capacity and self-induction and by elimination of C or of L reduces to the form encountered before. The impedance is $\sqrt{R^2 + \left(\dfrac{1}{pC} - Lp\right)^2}$ and the reactance is $\dfrac{1}{pC} - Lp$. The nature of the phase effect, whether lag or advance, depends on the relative magnitudes of Lp and $1/Cp$. If Lp is the greater $\tan \varphi$ is negative and φ is negative; that is, *the current lags behind the impressed E.M.F.* Otherwise there is a phase advance.

125. POWER FACTOR

The fact that we have phase changes in a circuit upon which an outside E.M.F. is applied when either inductance or capacity is present in the circuit leads to an important correction in the question of power

measurement. It will be remembered that power was defined as electromotive force times current. Thus the power in watts = volts $\times$ amperes.

In the case of alternating currents we measure virtual volts, that is,

$$E_v = \frac{E_0}{\sqrt{2}},$$

and we measure virtual amperes

$$i_v = \frac{i_0}{\sqrt{2}}.$$

The power in watts consumed would be therefore

$$E_v i_v = \frac{E_0 i_0}{2}.$$

These measurements could be made independently and the proper watts computed in the case where the current was not varying, or where there was no phase lag involved. Suppose, however, that instead of such conditions the E.M.F. of a circuit be 90° out of phase with the current flowing in it. This would be an extreme case but it is closely approached with high inductances and small capacities. The result would be that although the voltmeter correctly read the voltage and the ammeter correctly read the current the watts computed would be incorrect, for when the voltage was a maximum the current would be 0, and the actual power would be 0, while a finite amount was computed. This can best be understood by studying the equations. Suppose the instantaneous electromotive force is given by

$$E = E_0 \sin (pt + \varphi),$$

and the instantaneous current is given by

$$i = i_0 \sin pt.$$

The instantaneous power P will be

$$\begin{aligned} P &= [E_0 \sin (pt + \varphi)]\, i_0 \sin pt \\ &= E_0 i_0 \sin pt\, (\sin pt \cos \varphi + \cos pt \sin \varphi) \\ &= E_0 i_0 \sin^2 pt \cos \varphi + \tfrac{1}{2} E_0 i_0 \sin 2\, pt \sin \varphi. \end{aligned}$$

If we take the average value $\overline{P}$ of P, we must proceed as before and write

$$\overline{P} = \frac{\int_0^{\frac{\pi}{p}} P dt}{\int_0^{\frac{\pi}{p}} dt} = \frac{E_0 i_0}{\pi/p}\left[(-\tfrac{1}{2}\cos pt \sin pt + \tfrac{1}{2}pt)\cos\varphi - \tfrac{1}{2}\cos 2pt \sin\varphi\right]_0^{\frac{\pi}{p}}$$

$$= \frac{E_0 i_0}{2}\cos\varphi$$

and hence

$$\overline{P} = E_v i_v \cos\varphi$$

Consequently, the result will be that the average watts, or $\overline{P}$, is given by

$$\overline{P} = \tfrac{1}{2}E_0 i_0 \cos\varphi.$$

If $\varphi = 90°$, $\cos\varphi = 0$, and the average power consumption $\overline{P}$ is 0. If $\varphi = 0$, $\overline{P} = \frac{1}{2}E_0 i_0$. In between, $\cos\varphi$ has all values and the power consumed may have any value between

$$\tfrac{1}{2}E_0 i_0 = E_v i_v, \text{ and } 0.$$

Thus, it is seen that if we take the voltmeter and ammeter readings in the case of a circuit with self-induction and capacity we are in danger of not reading the true power consumption. What we read is called the *apparent watts*. The apparent watts times the $\cos\varphi$ gives the *true watts*. The *power factor*, as $\cos\varphi$ is called, gives the ratio of

$$\frac{\text{true watts}}{\text{apparent watts}} = \frac{\overline{P}}{E_v i_v} = \cos\varphi.$$

Hence the *power factor equals cos φ*. A wattmeter which has its fields excited by the current in the circuit, and its moving coil excited by the potential difference, will give the true power consumption because the phase differences in the two parts of the instrument give a simultaneous reading of the product of the true current and voltage combined in their right phase relationship. We can determine the power factor very easily by taking the wattmeter reading on a given circuit, and *then* taking the voltmeter-ammeter reading. The ratio of the wattmeter reading to the voltmeter-ammeter reading gives the power factor, and *so gives the cosine of the angle of phase lag*. We can consequently at once calculate the phase lag.

126. ELECTRICAL OSCILLATIONS

There is one more question which can only be touched on at this point and that is the ability of self-induction and capacity when united to give electrical oscillations. Historically, this field is exceedingly interesting. In Maxwell's mathematical formulation of Faraday's laws of magnetic and electric force there was the prediction that there must be in empty space a medium which Maxwell termed the luminiferous ether, capable of transmitting electromagnetic disturbances, which travel with the velocity of light. In fact, he deduced the fact that the ratio between the electromagnetic units (which depend on moving electricity), and electrostatic units, must be whole multiples of the velocity of light which is closely 3×10^{10} cm per second. This we have seen to be the case in a study of our units, see page 69. He furthermore predicted that it should be possible to generate electrical waves and to transmit them through the air. The oscillatory nature of a spark discharge under some circumstances was discovered and the equations in part worked out by Lord Kelvin. To Hertz belongs the credit of having utilized these oscillations for the transmission of electrical waves through empty space. This was the forerunner of the tremendous development which we have today in which the human voice can be transmitted across the continent, and in which we are now able to transmit from an aeroplane a picture of a field of battle and the progress of troops to headquarters 20 miles away by radio by a process popularly called television. The oscillatory nature of the electrical discharge is of interest to us inasmuch as it results from the combination of capacity and self-induction under the proper conditions.

Let us consider two tanks of water connected by a stopcock with a hole in it. If the stopcock has a small hole of large resistance and we have one tank full and the other empty, on opening the stopcock water will gradually flow from the full tank to the empty tank and when the two reach the same level the flow will stop. This is because the potential energy of the water which was initially higher than the final level was converted into kinetic energy in the pipe and this due to friction was converted to heat. Again, suppose the stopcock to have a large hole with very little friction. Then on opening the stopcock there will be a sudden rush of water from the full tank to the empty one. In the empty one the water will rise up to a height nearly as great as the height in the initial tank, owing to its inertia. It then rushes back into the initial tank, and its inertia carries it up to nearly its former position. Thus the water will oscillate back and

forth eventually coming to rest owing to the losses of kinetic energy in friction through the tube.

In electricity we have the almost analogous case. Suppose we have a capacity that is equivalent to our water tank. Assume we have the capacity charged but the circuit open. If we close the circuit the electricity will flow from one plate to the other, but it lacks one factor to make oscillation complete. The water beside having a capacity present had to have an inertia to carry it past its resting point. In electricity, the inertia is supplied by a self-induction. Thus, assume we have a capacity connected across a self-induction as indicated in Fig. 126. If now the capacity C be charged up and the switch S closed the positive electricity will flow through the inductance L to the other side of the circuit. Owing to its inertia it will charge the other side of the condenser to a higher potential than the equilibrium position of its charge. The electricity will then flow back again and oscillation will continue until the resistance damps out all oscillations. If we have a high resistance we will have the same case we had in the two tanks of water with a small hole, that is, a gradual flow and no oscillation. The condition for oscillation then is that the capacity and self-induction be sufficient, but that the resistance be small enough relative to the two to prevent damping out the oscillations. Again, it is perfectly obvious from our water analogy that the period of the oscillation will depend on the capacity and self-induction present. As a matter of fact the period is

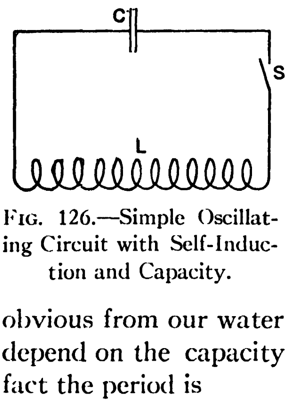


FIG. 126.—Simple Oscillating Circuit with Self-Induction and Capacity.

$$T = 2\pi\sqrt{LC},$$

to a good degree of approximation. A further study of such oscillating circuits belongs in a more advanced course. The study of such oscillations is absolutely essential to all investigation of radio engineering.

127. DERIVATION OF THE EQUATION FOR OSCILLATORY DISCHARGE OF A CONDENSER WITH RESISTANCE CAPACITY AND SELF-INDUCTION

For those students who wish to delve into the nature of electrical oscillations a little more deeply than the time in such a course as this permits, the equations for the discharge of a condenser of capacity C through a self-induction L and a resistance R are given in a compact

form below. Consider the circuit shown in Fig. 127 containing C, L, and R as indicated. Assume the condenser C charged to a potential E by closing the switch K to the battery. Suppose that now the switch K is opened, and the switch K' to the self-induction and capacity, which was previously open, is closed. Current then starts to flow through the circuit and we can quantitatively consider the conditions which follow as outlined qualitatively in the preceding section. In the mesh CRL the E.M.F. applied is now zero. At any instant, however, we have an E.M.F. E'' due to the charge on the condenser, and E.M.F. E' of self-induction and the iR drop in the line. Had the E.M.F. applied been E the complete Kirchhoff law equation for the mesh would have been $E - (E' + E'') = Ri$ as indicated in the earlier part of this chapter for the case of an alternating current E applied to a circuit with self-induction, resistance and capacity.

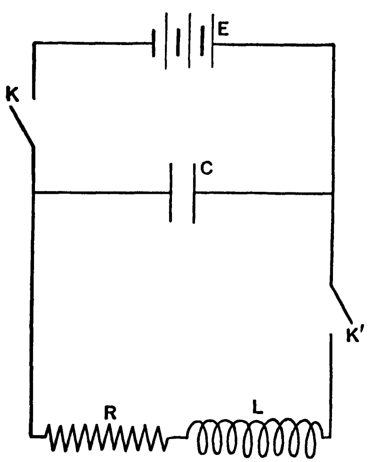


FIG. 127.—Circuit for the Study of the Discharge of a Condenser Through an Inductance and a Resistance.

Thus $E = Ri + E' + E''$, where in the present case $E = 0$, $E' = \dfrac{Q}{C}$ and $E'' = L\dfrac{di}{dt}$. As $i = \dfrac{dQ}{dt}$, the equation for our simple system becomes

$$L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{Q}{C} = 0.$$

This is called a linear differential equation of the second order and first degree in Q with constant coefficients. For convenience we may write

$$\frac{d^2Q}{dt^2} + \frac{R}{L}\frac{dQ}{dt} + \frac{Q}{CL} = 0.$$

Such an equation may be solved by setting $Q = e^{\alpha t}$ as a solution and studying the effect on the equation. This treatment yields

$$\alpha^2 e^{\alpha t} + \frac{R}{L}\alpha e^{\alpha t} + \frac{1}{CL}e^{\alpha t} = 0,$$

whence

$$\alpha^2 + \frac{R\alpha}{L} + \frac{1}{CL} = 0.$$

This is a second degree equation in the constant α whose algebraic solution takes the form

$$\alpha = -\frac{R}{2L} \pm \frac{\sqrt{\frac{R^2}{L^2} - \frac{4}{CL}}}{2}.$$

This gives us at once two solutions

$$Q = A'e^{\left(-\frac{R}{2L} + \frac{1}{2}\sqrt{\frac{R^2}{L^2} - \frac{4}{CL}}\right)t}$$

$$Q = B'e^{\left(-\frac{R}{2L} - \frac{1}{2}\sqrt{\frac{R^2}{L^2} - \frac{4}{CL}}\right)t}.$$

Here A' and B' are arbitrary constants of integration to be evaluated later. For convenience in handling one can call $\frac{R}{L} = 2b$ and $\frac{1}{LC} = k^2$. The equations become

$$Q = A'e^{(-b + \sqrt{b^2 - k^2})t}$$

and

$$Q = B'e^{(-b - \sqrt{b^2 - k^2})t}.$$

As

$$Q - A'e^{(-b + \sqrt{b^2 - k^2})t} = 0$$

and

$$Q - B'e^{(-b - \sqrt{b^2 - k^2})t} = 0$$

we can write the complete solution as

$$2Q - A'e^{(-b + \sqrt{b^2 - k^2})t} - B'e^{(-b - \sqrt{b^2 - k^2})t} = 0,$$

and calling $\frac{A'}{2} = A$, and $\frac{B'}{2} = B$, we have

$$Q = Ae^{(-b + \sqrt{b^2 - k^2})t} + Be^{(-b - \sqrt{b^2 - k^2})t}.$$

Now we can find A and B. For since $Q = Q_0 = EC$ at $t = 0$

$$A + B = Q_0 = EC.$$

Again, since when $t = 0$, $\frac{dQ}{dt} = i = 0$, we have

$$(-b + \sqrt{b^2 - k^2})A + (-b - \sqrt{b^2 - k^2})B = 0.$$

These two equations for A and B may now be solved, and we obtain for Q the equation

$$Q = Q_0\left\{\frac{1}{2}\left(1 + \frac{b}{\sqrt{b^2 - k^2}}\right)e^{(-b + \sqrt{b^2 - k^2})t} + \frac{1}{2}\left(1 - \frac{b}{\sqrt{b^2 - k^2}}\right)e^{(-b - \sqrt{b^2 - k^2})t}\right\}$$

It is seen that Q at any time t is a function of $Q_0 = EC$ and the constants $b = \frac{R}{2L}$ and $k^2 = \frac{1}{CL}$. When $b^2 - k^2$ is positive the exponents are positive and we have the case of a decline in Q which is the initial charge on the condenser multiplied by the sum of two exponential terms which decrease as t increases. The quantity and hence the current which is $\frac{dQ}{dt}$ thus decreases exponentially with time. This condition of b^2 greater than k^2 corresponds to a high resistance compared to capacity, for unless $b^2 = \frac{R^2}{4L^2}$ is greater than $k^2 = \frac{1}{CL}$, the condition above does not hold. Hence for values of C, R and L such that $\frac{R^2}{4L^2}$ is greater than $\frac{1}{CL}$ the flow of current is the exponential decrease characteristic of a water connection with high resistance where the energy of current flow is consumed in heating the tube as it flows through.

If now b is less than k, or $\frac{R^2}{4L^2}$ is less than $\frac{1}{CL}$, the case is entirely different. The two exponential terms as well as the quantities A and B become imaginary, that is, they contain the square root of -1. Thus $b^2 - k^2$ is now negative, for k^2 is greater than b^2, and we should write $\sqrt{-1}\,\sqrt{k^2 - b^2}$. The $\sqrt{k^2 - b^2}$ is then real and we can represent the $\sqrt{-1}$ by the symbol j. In writing the equations for this case the parts of the terms

$$e^{-(b \pm \sqrt{b^2 - k^2})t}$$

can be written in the form

$$e^{-bt}\, e^{\pm(\sqrt{b^2 - k^2})t}.$$

The equation deduced for Q then becomes, if we collect the terms multiplied by j and by $\frac{b}{\sqrt{b^2 - k^2}}$ separately,

$$Q = Q_0 e^{-bt}\left\{\frac{e^{j\sqrt{k^2 - b^2}t} + e^{-j\sqrt{k^2 - b^2}t}}{2} + \frac{b}{j\sqrt{k^2 - b^2}}\,\frac{e^{j\sqrt{k^2 - b^2}t} - e^{-j\sqrt{k^2 - b^2}t}}{2}\right\}$$

Now by Euler's theorem the exponentials $\dfrac{e^{jat}+e^{-jat}}{2} = \cos at$, and $\dfrac{e^{jat}-e^{-jat}}{2j} = \sin at$. If $a = \sqrt{k^2 - b^2}$ then the above equation for Q at once simplifies to the form

$$Q = Q_0 e^{-bt}\left\{\cos\sqrt{k^2 - b^2}\,t + \frac{b}{\sqrt{k^2 - b^2}}\sin\sqrt{k^2 - b^2}\,t\right\}$$

and by a well known theorem in trigonometry similar to that used in deducing the current relations on page 292,

$$Q = \frac{Q_0 k e^{-bt}}{\sqrt{k^2 - b^2}}\cos\left(\sqrt{k^2 - b^2}\,t - \varphi\right),$$

where

$$\tan\varphi = \frac{b}{\sqrt{k^2 - b^2}}.$$

Thus we can write in terms of R, L and C

$$Q = Q_0 \frac{e^{-\frac{R}{2L}t}}{\sqrt{LC}\sqrt{\dfrac{1}{LC} - \dfrac{R^2}{4L^2}}}\cos\left(\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}\,t - \varphi\right)$$

and

$$\tan\varphi = \frac{R}{2L\sqrt{\dfrac{1}{CL} - \dfrac{R^2}{4L^2}}}.$$

We accordingly see that we have: (1) a discharge which varies periodically with time, for $\cos\sqrt{k^2 - b^2}\,t$ varies periodically with time; in other words when k^2 is greater than b^2 we have an oscillatory discharge; (2) the oscillation is one of decreasing amplitude; that is, it is a damped oscillation, for the cosine term is multiplied by the exponential $e^{-\frac{R}{2L}t}$ which decreases with time, and this represents a damping factor (see Chapter XXIII); and (3) we have a phase factor φ introduced into the equation where φ can be evaluated by

$$\tan\varphi = \frac{b}{\sqrt{k^2 - b^2}}.$$

The period of the oscillation is determined by the coefficient of the t in the cosine term. Since by definition of an oscillatory phenomenon

(see Chapter XX) the period T is given by the T in the transcendental term, sine or cosine, in the form $\sin \frac{2\pi}{T} t$, we see at once that

$$\frac{2\pi}{T} = \sqrt{k^2 - b^2}$$

or

$$T = \frac{2\pi}{\sqrt{k^2 - b^2}} = \frac{2\pi}{\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}}.$$

Thus the oscillatory discharge will have a damped oscillation of period T defined by

$$T = \frac{2\pi}{\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}}.$$

Now where one desires sustained oscillations one reduces R to a minimum and the term $\frac{R^2}{4L^2}$ becomes negligible. In this case the terms containing $\frac{R^2}{4L^2}$ are simplified by putting $\frac{R^2}{4L^2} = 0$ and the elaborate equation for Q above becomes $Q = Q_0 \cos \frac{t}{\sqrt{LC}}$, for φ is 0 and $e^{-\frac{R}{2L}t}$ becomes unity. Hence for this case the damping is negligible, and the period is $\frac{2\pi}{T} = \frac{1}{\sqrt{LC}}$ or $T = 2\pi\sqrt{LC}$ as stated in the simplified equation given in the preceding qualitative discussion.

CHAPTER XXVI

DISCHARGE THROUGH GASES AND ATOMIC STRUCTURE

128. DISCOVERY OF X-RAYS

IN 1895 a German physicist, W. C. Roentgen, was studying the electrical discharge of an induction coil in an evacuated tube made possible by the advance in the development of vacuum technique. He noticed a peculiar greenish fluorescence when the vacuum in the tube became low enough. This greenish fluorescence was accompanied by a stream of bluish rays moving in straight lines from the negative electrode. By accident, he observed that a screen of a material called barium platino cyanide lying on the table near the apparatus became phosphorescent whenever the greenish fluorescence appeared. Investigation showed that this was due to a new type of radiation which passed through the glass walls of the vessel and through air. The remarkable properties of these radiations, which were called x-rays because of their unknown origin, led to feverish experimentation in this field of work. These rays were found to cause fluorescence in certain substances, they were found to affect a photographic plate, and they were found to make the air through which they passed conducting. It was the conductivity of air, which had been considered almost a perfect insulator from the time of Coulomb, that enabled further fundamental researches in the study of x-rays and carriers of electricity to be made. The x-rays were found to travel in a straight line from the source and to cast shadows when thick opaque objects were placed in their path.

129. CATHODE RAYS

Immediately following the discovery of the x-rays came the investigation of the peculiar bluish streamers emitted from the negative electrode in the evacuated tube whose presence simultaneously with the x-rays was always observed and to whose presence the x-rays were ascribed. It was very quickly found by different observers that these rays were apparently negatively charged; that they were deflected in a magnetic field as if they were a current of negative electricity;

that they moved in straight lines; and that shadows cast by objects placed in the tube showed that where they did not strike the tube there was no fluorescence due to the x-rays. It was J. J. Thomson who succeeded in showing that these streamers were negative electrical particles moving with velocities of around one hundred million centimeters per second and having a ratio of electrical charge to mass equivalent to 1860 times that of hydrogen ions in electrolysis. The method by which he succeeded in proving this was that by putting a pair of electrodes into the tube he was able to deflect the stream with an electrical field. By combining the electrical and magnetic fields, using the following theoretical relations, J. J. Thomson was able to evaluate the velocity v and the ratio of $\frac{e}{m}$, that of charge to mass, for these particles. In Fig. 128 we have a beam of cathode rays emanating from the cathode C defined by a slit S in a plate D, which is the anode, and through which the rays pass because of their velocity gained in the evacuated region between C and S. In the absence of an electrical field these rays would pass between the two sealed-in electrodes A and B constituting a parallel plate condenser and impinge in a point on the fluorescent zinc sulfide screen P. When a field is placed on the plates A and B as indicated in the picture the electrons are deflected downward and the following relations may be deduced. If the rays are particles moving with a velocity v and each carries a charge e the force acting at right angles to the electrons in an electrical field of X volts per cm would be Xe. Thus as long as the particles remain between the plates they are urged towards the positive plate with a force $f = Xe$, and an acceleration $\frac{f}{m} = \frac{Xe}{m}$, where m is their mass. They are thus, just as is a stone that travels with a horizontal velocity v and is urged to the earth with a constant force mg, forced to describe a parabolic path. If the particles while moving the length d of the plates were deflected downward by the field a distance x, as shown in Fig. 128, the same calculations as apply to the stone could be applied to this case. In a time t they move the length

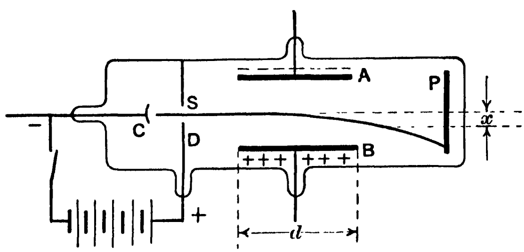


FIG. 128.—Determination of $\frac{e}{m}$ for Cathode Rays, Electrostatic Deflection.

of the plates d, so that $d = vt$. Now in the same time t the particles moved downward a distance x. However $x = \frac{1}{2}at^2$, where a is the acceleration, so that $t = \sqrt{\frac{2x}{a}}$, where a the acceleration in this case is $\frac{Xe}{m}$. Hence $d = vt = v\sqrt{\frac{2x}{\frac{Xe}{m}}}$ and we can write

$$\frac{d^2}{v^2} = \frac{2x}{X\frac{e}{m}},$$

or

$$\frac{e}{m} = \frac{v^2}{d^2}\frac{2x}{X}.$$

This equation has three unknowns in it: e the charge of the particle, m its mass, and v the velocity, and we can obtain *only* the values of $\frac{e}{m}$ and v from this, and that only with the aid of another experiment. If we let a magnetic field H act on the stream of these particles so that H is perpendicular to the plane of Fig. 129, the beam of particles experiences a force constantly at right angles to their motion. Now, as we will recall, such a force causes a circular path to be followed for which the equation for centripetal force $\frac{mv^2}{\rho} = f$ holds, where m is the mass of the particles, v is their velocity and ρ is the radius of curvature of their path. ρ can of course be measured by the deflection x of the beam along a distance d. Now a current i encounters a force Hi per unit length in a magnetic field, and a particle of charge e with a velocity v constitutes a current i. Thus $f = Hev$. Accordingly we may write that

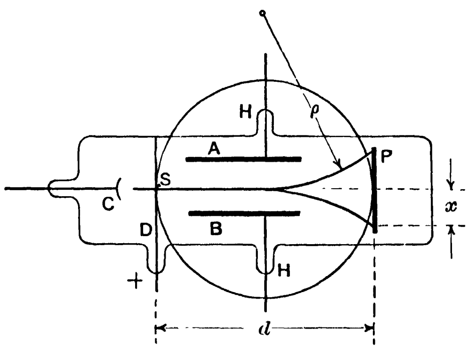


FIG. 129.—Determination of $\frac{e}{m}$ for Cathode Rays, Magnetic Deflection.

$$Hev = \frac{mv^2}{\rho}, \text{ whence } \frac{e}{m} = \frac{v}{H\rho}.$$

Thus by measuring the ρ for a given H in one experiment and by measuring x and d for the same rays in a field X in a second experiment we have two equations containing the unknown quantities $\frac{e}{m}$ and v. To get e and m separately requires still another equation in e and m and the same v. Such an experiment could not be devised at that time. J. J. Thomson found for $\frac{e}{m}$ a value for these rays which was 1860 times the value found for the hydrogen ion in electrolysis, and the quantity $\frac{e}{m}$ was negative as these rays carry a charge of negative sign. The velocity v found depended on the electrical potential P.D. across the electrodes C and D and was of the order of 10^8 cm/sec.

Later when it was found possible to generate these electrical particles independently of the gaseous discharge a very simple set of relations for finding $\frac{e}{m}$ was devised. Since a particle of charge e freely acted on by an electrical potential difference P.D. acquires a kinetic energy of motion in the absence of a gas, it is at once clear that the kinetic energy of the particles $\frac{1}{2}mv^2$ must equal PDe, the work done on the particle by the field. Hence

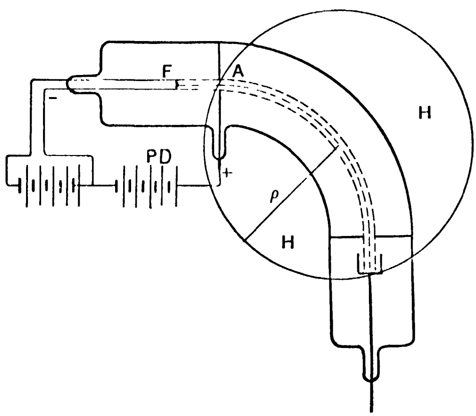


FIG. 130.—Determination of $\frac{e}{m}$ for Cathode Rays from a Hot Filament Using the Energy of the Electrons Acquired from the P.D. and Magnetic Deflection.

$$PDe = \tfrac{1}{2}mv^2 \text{ and } \frac{e}{m} = \frac{v^2}{2PD}.$$

Now by the arrangement of Fig. 130 the particles from a hot negative filament F in a high vacuum are driven by a potential difference PD between a plate A which is positive and the filament F through a hole in A. Since the particles move very rapidly many of them pass through the hole without being caught by A. This stream of particles on the other side of A enters a magnetic field H perpendicular to the paper and into it, which forces the beam in the circular path ρ. As ρ can be measured and as H and PD are known the relations $\frac{e}{m} = \frac{v^2}{2PD}$ and

$\frac{e}{m} = \frac{v}{H\rho}$ enable $\frac{e}{m}$ to be determined. The experiments showed these streamers to be made of negative particle of velocities dependent on the P.D., of $\frac{e}{m} = 5.303 \times 10^{17}$ E.S.U. per gram. Subsequent experiments of Millikan and others showed that the particles had a negative charge $e = 4.77 \times 10^{-10}$ E.S.U. and a mass $m = 8.99 \times 10^{-28}$ grams. The particle was called the electron. Subsequent reasearch showed that it was the mobile part of electricity in all solid bodies, and that electrification was caused by the removal of electrons from bodies, leaving them positively charged. It was further found that all matter contained electrons, and J. J. Thomson postulated the *electron theory of matter.*

130. CONDUCTION IN GASES—GASEOUS IONS

The great conductivity of gases produced by the x-rays led to a study of the mechanics of conduction of electricity in gases. Previous to this, Coulomb had proven that the air conducted electricity to a very slight amount. The amount was so small, however, that he was unable to study the mechanism. It is now known that what the x-ray does, and what is often done by light rays, is to tear rapidly moving electrons from the molecules of the walls of the vessel or of the gas, and to send them crashing through many thousands of atoms. Every once in a while such an electron strikes an atom so as to tear more electrons out. We then have atoms which have lost an electron, and they have a positive charge. Some of the electrons after losing their speed finally attach to molecules and make negative gaseous ions. If we pass x-rays through a gas with two plates immersed in the gas and place a potential difference between the plates a current will pass. In general, the current is so small that we must use electrometers, or electroscopes, and small capacities to measure them, for the number of molecules ionized is small and the charge they carry is only 4.8×10^{-10} E.S.U. each. It would be observed that as we increase the potential between the plates the current increases at first linearly, then more slowly, and finally reaches a saturation or constant value. When this occurs we state that we have saturation and we assume that what has happened is that the field is removing all of the ions that are formed as fast they are formed. Before this point was reached, the field was so weak that before the ions got across some of the ions of opposite sign had recombined. These attract each other strongly according to an inverse square law of force. That is, the

negative ion with its extra electron had joined a positive ion and given up its electron to it. Such currents imply that the ions in a field must move with a finite velocity. We have succeeded in measuring in many ways the speed with which these ions move in gases. The value of the velocity in air is approximately 2.2 cm per second in unit electrical field for the negative ions, and 1.6 cm per second in unit electrical field for the positive ions. The velocity of the ion in unit electric field is called mobility and it is inversely proportional to the gaseous pressure through large ranges of pressure. In certain pure gases like He, Ar, N_2 and H_2 the electron remains free for indefinite periods. In other gases, especially the more chemically active electronegative gases, the electron may attach rather readily. Its ability to attach varies widely. In some cases attachment takes place at the first impact, while in others the electron may make 10^6 impacts before it can strike a molecule in such a fashion as to attach and form a negative ion. The electron in air as long as it remains free moves with a speed of the order of 10,000 cms a second in unit field.

If the field in which an ionized gas is placed be increased beyond the value of the saturation current, the current will suddenly increase sharply, eventually being followed by a spark if the field is high enough. This is ascribed to the action of the electrons and positive ions in moving so fast in the high field that they knock new electrons out of previously neutral molecules, and so generate enough ions to carry the current. When the high fields produce this ionization of molecules or atoms by collision under conditions where the energy supplied to the electrode is great enough to maintain the fields the whole gas becomes highly conducting through the accumulation of ions and electrons and we have a gaseous discharge which, depending on conditions of voltage, gas pressure and electrode characteristics, gives either a spark, an arc, or a glow discharge such as is seen in the modern neon advertising sign. Where the gas atom or molecule has lost an electron and regains that electron by some recombination process as is the case in the luminous discharges just mentioned the electrons on recombining with the ions emit a radiation which gives some of the light observed in discharge tubes. In other cases, the electron on striking an atom is unable to remove an electron but can raise it to a position in the atom of higher potential energy. In most gases it may remain in this state some 10^{-8} seconds, after which it returns to its initial state and emits one of the characteristic lines of the spectrum of the atom disturbed. In certain atoms, e.g., mercury and neon, this disturbed state lasts for as long as 10^{-4} seconds. Such states are called metastable. They enable ionization to occur in suc-

cessive stages and are probably responsible for the efficient operation of low voltage arcs such as a mercury arc.

131. THE ELECTRON

Investigation has shown that the value of Ne, the Faraday constant, for the gas ions, is the same as for univalent negative ions in electrolysis. The conclusion is that the ions are of the order of magnitude of molecules with a single electrical charge, there being perhaps several molecules to the ion. Millikan, following the lead of C. T. R. Wilson, J. J. Thomson, and others, succeeded in catching single ions on oil drops. He balanced the oil drop with its charge in an electrical field against the earth's gravitational field. The result led him to a measurement of the value of the elementary negative charge on the electron, and he found that it was 4.77×10^{-10} electrostatic units of electricity. Thus we see that our electrostatic unit of electricity is a very large multiple of the real elementary charge, the electron. From this determination and the measurement of $\frac{e}{m}$ the mass of the electron was shown to be $\frac{1}{1860}$ that of the hydrogen atom; i.e., the mass of the electron is about 8.9×10^{-28} grams. Its diameter has never been accurately measured, but the deflection experiments of Sir Ernest Rutherford, to be mentioned later, showed that it must have dimensions of the order of magnitude of 5×10^{-13} cm.

132. RADIOACTIVITY

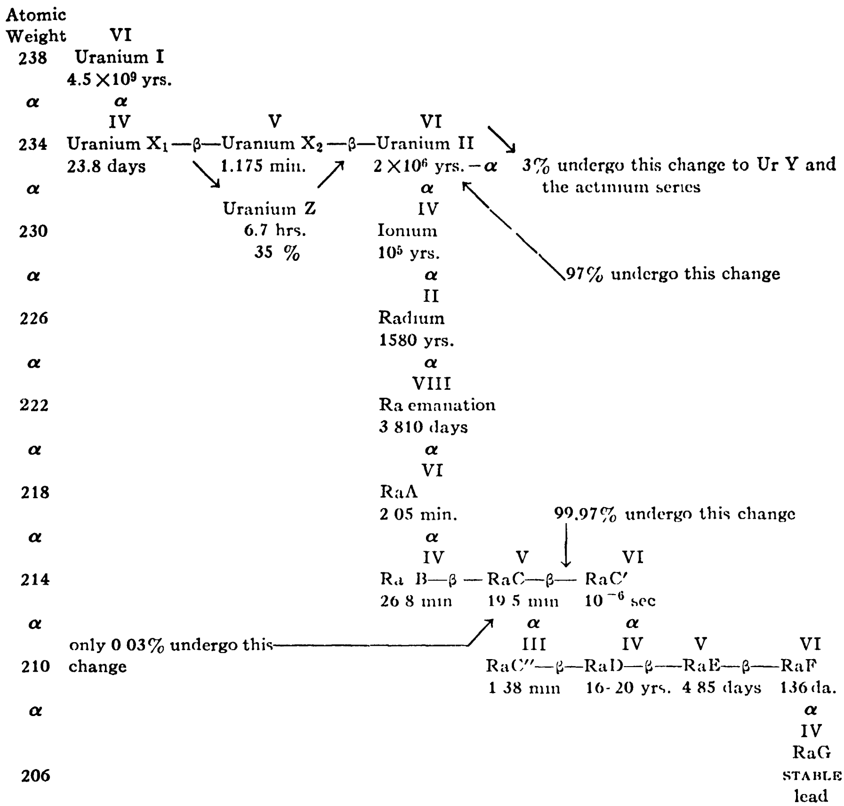
Shortly after the fluorescence discovered by Roentgen as a result of the action of x-rays, Becquerel studied fluorescent substances which had no relation to x-rays to see whether such substances would affect a photographic plate. Among others, he tried some beautiful fluorescent uranium salts. Exposure in a dark room for several days showed this substance to affect a photographic plate. Investigation of many chemical elements immediately followed, and it was found that thorium was even more active on the photographic plate than uranium. Investigation quickly showed that thorium produced an increased conductivity in the air about it, and thus the phenomenon of radioactivity was discovered. Madame Curie, with the skilful cooperation of her husband, isolated from certain pitchblende ores a substance thousands of time more efficient in these activities than thorium. It was called radium. Under the genius of Sir Ernest Rutherford, Soddy, Madame Curie, and others, the isolation of many radioactive products was achieved. Rutherford showed that one

radioactive atomic substance changed into another, giving out radiations which affected the photographic plate and ionized the air. It was shown that uranium was the great-great-great-grandparent of radium. These radiations were shown to be of three types: *alpha rays*, which Rutherford, with the greatest ingenuity, proved beyond doubt to be atoms of helium gas charged with two positive charges moving with a velocity from 2×10^9 cm/sec to 2×10^8 cm per second; *beta rays*, which were shown to be negatively electrified particles, called electrons, moving with velocities in some cases 99 per cent that of light; finally there were radiations of the nature of x-rays which penetrated several inches of lead—these were called the *gamma rays*. Each radioactive change is accompanied by an emission from the inner part or nucleus of the atom of either an α ray or a β ray. The γ rays rise from most β ray changes and are also observed in a weak form in some α ray changes. The relation of γ rays to the process is not clearly understood. It is known that the β ray change precedes the γ or α ray. The γ rays from α particle changes are in most cases soft x-rays resulting from α ray impact.

The Disintegration Series of Uranium.—The table below illustrates a radio active disintegration series and represents the disintegration series of the uranium-radium family. A displacement downwards in a column indicates a loss of one α-particle, i.e., a decrease of 4 in atomic weight and change of charge by loss of 2+ charges. A displacement to the right in a row indicates the loss of one β particle, i.e., the loss of one electron, but no change in atomic weight and an increase of + charge by one unit. The Roman numerals above each member indicate the group to which it belongs in the periodic table. The time periods below each member indicate its "half-life" period, T, which is related to its decay constant λ defined by $N = N_0e^{-\lambda t}$, by $T = \dfrac{0.69}{\lambda}$.

133. STRUCTURE OF THE ATOM

It was found by Rutherford, and later Fajans, Soddy, and others, that when these radioactive series of changes took place an element that lost an α particle changed its chemical properties in such a way that it moved two places to the left in the periodic table of elements. That is, radium, an element like barium, loses an α particle with two positive charges from its nucleus, and shifts into a gas which is chemically inert and is called niton. It falls into the same class of elements with helium and argon and decreases its mass by 4 units of atomic weight in this change; and similarly polonium, which falls into the sixth group in the Periodic Table (i.e., is in the same group with oxygen),



loses an α particle and changes to *RaG*, which is identical in properties with lead. For those changes in which a fast electron comes from the nucleus the atom does not change in weight, except by the mass of the electron, but it moves one place to the right in the Periodic Table. Thus *RaD* which is similar to lead in the fourth group of the Periodic Table loses a β ray and gives *RaE*, which is similar in its properties to bismuth, belonging to the fifth group of the Periodic Table, and again *RaE* loses an electron and changes to *RaF* or polonium, which belongs in the sixth group of the Periodic Table. The mass of these atoms is all the same except for the loss of two electrons whose mass is negligible. We thus see that the atomic weight is not the significant thing in the chemical behavior of the elements, while the nuclear charge or the charge on the internal part of the atom is the determining factor. The reason for this will become obvious later. Inasmuch as beside the radium-uranium series of radioactive elements there are two other series, the thorium series and the actinium series, all of them having atomic weights in this region of the Periodic Table and

giving many products of α and β changes, there must be several of the positions in the Periodic Table which are simultaneously occupied by the radioactive elements or their end products which have similar chemical properties and inner charges, while their atomic weights may differ by considerable amounts. It is now known that ordinary lead, whose atomic weight is determined as 107, is composed of at least two leads, the ultimate results of disintegration of radium and thorium, having atomic weights 106 and 108, respectively. Such elements with similar chemical properties and occupying the same position in the Periodic Table but having different atomic weights are called *isotopic* elements. The existence of these was first discovered by J. J. Thomson in the study of positive rays to be discussed. He discovered that neon gas consisted of two elements with identical physical and chemical properties but different atomic weights in the neighborhood of 20 and 22. The development of Thomson's method by Aston and Dempster has led to the discovery that almost all of the elements in the Periodic Table are isotopic. The atomic weights are therefore very closely whole numbers, and where we find elements having fractional atomic weights like chlorine we know that they are composed of two or more isotopic elements having different atomic weights (in this case, 37 and 35, respectively). Thus these investigations have shown us that matter is made of electrons and positive charges some of which seem to come out of the internal structure of the atom (nucleus) in the form of β and α particles, and it is seen that the number of these charges in the internal structure of the atom profoundly determine the chemical properties, while the weight of the atom is of distinctly less significance.

While all this was going on, a man by the name of Goldstein observed in a cathode ray tube that there was projected through a hole in the cathode a beam of rays which had a positive charge moving in the opposite sense from the cathode rays. These were studied more accurately by W. Wien, and finally by J. J. Thomson, who showed that the value of $\frac{e}{m}$ for these particles indicated them to be positively charged atoms of matter that had lost electrons. In hydrogen he was never able to knock out more than one electron; in lithium no more than three. In heavier elements, like Hg, he was able to knock out as many as eight, which was the highest number observed, and was observed in the case of mercury. This led to the conclusion that the positive charge always remains with the atom while electrons may be knocked off. It was Rutherford who showed by the scattering of α particles by atoms that the electrons which made the atom neutral

were moving in the outer part of the atom while the positive electricity was contained in a minute *nucleus* in the center of the atom. The size of this nucleus was of the order of magnitude of 10^{-12} cm while the diameter of the atom was of the order of 10^{-8} cm and the force between α particles and the nucleii of the atoms closely follow Coulomb's law of force down to distances of the order of magnitude of 10^{-12} cm. Rutherford and his pupils were also able to measure the charge on the nucleus and found that the charge on the nucleus corresponded very closely to a law deduced by Moseley, which said that *the charge on the nucleus was equal to the number of the element in the Periodic Table of the elements.* This explains why hydrogen, the first element, had only one external electron, for it has only one positive charge. Helium has only two external electrons because it has only two positive charges. Sodium has eleven external electrons because it is the eleventh element in the Periodic Table, and has eleven units of positive electricity on its nucleus. Rutherford was later able to show that by shooting fast helium atoms, or α particles, into the nucleii of atoms, like nitrogen, aluminum, and others, he was able to knock out hydrogen nucleii or protons, of very high velocity. In doing this the atom of nitrogen was changed chemically, as it had its nuclear charge altered, making it a new element. The α particle in these encounters as visually shown by Blackett goes into the nucleus with its mass of 4 units and its charge of $2+$ units, and an H nucleus with mass 1 and $1+$ charge comes out; the new element has a charge of $8+$ and is similar to oxygen with an atomic weight of $14 - 1 + 4$, or 17. Thus we observe here an actual case of the transmutation of an element caused by an external source. The radioactive elements are elements which are spontaneously undergoing transmutation. The artificial transmutation here described however is exceedingly inefficient. One must shoot on the average ten million fast α particles into a gas like nitrogen to get one atom of nitrogen transmuted to its oxygen-like product. It is therefore improbable that we have achieved the dream of the old alchemists of being able to change lead into gold cheaply, inasmuch as the energy expenditure involved is worth millions of times the value of the gold. Hence, observationally, from radioactive and artificial disintegrations we find the *nucleii* of the elements to consist of helium nucleii and hydrogen nucleii with some electrons as the β rays indicate. This enables us to picture the constitution of the nucleii. In an element, for instance, like helium which has a mass 4, and a charge 2, there must be 4 hydrogen nucleii to give it its mass, and 2 electrons to give it a charge 2. The hydrogen nucleus, which is the *smallest positive charge* and whose charge is equal

to that of the electron but opposite in sign, is the unit of positive electricity and is called the proton. It is much smaller than the electron, though more massive, its effective radius being possibly of the order of 10^{-16} cm.

We see thus that matter is made up of positive protons and negative electrons. It is further seen that there are two types of structure into which these units can combine in building up the atoms of matter. Electrons and protons combine together in very compact forms with a residual positive charge giving the atomic nucleii. These are small and are of the order of 10^{-12} cm in diameter or less. The net number of positive charges on any given nucleus determines furthermore the number of electrons which are held in orbits about a neutral atom. Thus an atom with a positive nuclear charge of 20 units having a nucleus made up of 10 α particles or 40 protons, and 20 cementing electrons, will have moving about its *minute nucleus* twenty other electrons moving in orbits whose average radii fall into different groups. In a neutral atom, however, as electrons are added to the outer structure of the atom, they will arrange themselves in the most stable configurations. The work of Lewis, Kossel, and Bohr indicates that the preferential stable configurations are systems of 2, 8, 18, or 32 electrons in orbits. As the number of electrons exceeds one of these configurations added electrons go to the outside of the atom, starting a new group which is therefore dynamically unstable and will react violently chemically to attempt to complete its more stable configuration where possible. It will often do this at the expense of its electrical neutrality. (See Chapter XII.) Needless to say, where the incompleteness of the shell reaches dimensions of 16 out of 32 electrons, as in the longest period of the table, the change of one or more electrons will not improve conditions very much, and we find in this part of the Periodic Table rather indefinite chemical affinities, e.g., the rare earths.

In argon, the outer atomic electrons will first consist of two electrons fairly close to the nucleus, that is, about 10^{-10} cm away, then there will be a group of orbits having an average radius of the order of 10^{-9} cm, and finally there will be still another group of 8 electrons reaching well out to 10^{-8} cm, and which form the surface of the atom. Chemical combinations, such as we know them, depend entirely on the electrical forces and the affinities between the outer shells or groups of electrons and the rest of the atoms. They are consequently of a very superficial nature, and the distance being great in these cases the forces will be small compared to those inside the nucleii or in the inner shells. We thus see how a further study of the nature of

matter has ultimately led us back to a description of all matter based on electrical charges. It is, therefore, very necessary that we understand the laws of electricity in order to be able to further our knowledge of atomic nature.

134. ELECTROMAGNETIC RADIATIONS

A word or two more might be said about the so-called gamma rays and x-rays. X-rays and gamma rays have been shown to be electromagnetic radiations exactly similar in nature to light waves and the long Hertzian oscillations now used in radio. The only difference is one of wave length. The x-rays range in wave length from 10^{-7} cm to 10^{-9} cm as compared with light waves of 10^{-5} cm. Gamma rays are still shorter x-rays which come from radioactive transformations. They have wave lengths of from 10^{-9} to 10^{-11} cm. The gamma rays are different only in that they come from vibrations of electrons *inside the nucleus*, whereas the *x-rays mainly come from vibrations of the first two inner layers of electrons in the atom.* These are the layers at about 10^{-10} and 10^{-9} cm from the nucleus, and are called the K and L shells. The *visible and ultraviolet light waves* come from vibrations of electrons on the outsides of the atoms. Their wave lengths extend from 10^{-7} cm to about 10^{-4} cm. The long heat waves with which we are familiar come from the vibrations of *whole charged atoms in solids.* Their wave lengths extend to close to 10^{-2} cm. Beyond this we reach into the regime of short electromagnetic waves. For such waves, to get high frequencies, T must be small, and this means self-induction and capacity must be small. Such oscillations can be obtained from minute iron filings in oil when discharge conditions are proper. From those on to the infinitely long wireless waves is merely a question of the proper conditions of capacity and self-induction. It was to the credit of Laue, and of Bragg following him, that we were able to prove that x-rays are light waves and to measure their wave lengths. All our diffraction gratings have distances between the rulings too great to diffract light of the shortness in wave length of x-rays. Laue remembered that regular atomic spacing in crystals gave distances of the order of magnitude of the wave length of x-rays. He consequently concluded that the atoms in a crystal must be able to scatter x-rays and cause diffraction patterns. He showed this was the case. Today, we have even succeeded in *refracting* x-rays and measuring the index of refraction for them, as well as measuring the wave length directly with ruled diffraction gratings at grazing incidence, thus showing them to be in all respects identical in nature with the light waves. It is seen that these electromagnetic waves cover an

enormous range from kilometers in length to 10^{-11} cm. The only difference in them is the nature of their origin. An electrical body of small inertia, an electron tightly bound in the nucleus, gives the high frequency oscillations of γ rays. An inner electron in the neighborhood of the intense field of the nucleus will give a frequency characteristic of the x-rays. A valence electron or a lightly bound atom in a crystal or the molecule of a gas gives light or heat waves, and a mass of electrons surging in a system with inductance and capacity at a low frequency gives wireless waves.

Wave Length	Range	Source	Name	How and by whom studied
∞	to 100 cm	Movements of electricity in large systems with capacity and self-induction	Radio waves	Predicted by Maxwell Discovered by Hertz
100 cm	to .01 cm	Electrical oscillations of minute systems, metal filings.	Very short electromagnetic waves	Nichols and Tear Large gratings.
.01 cm.	to 7×10^{-5} cm	Oscillations or vibrations of charged atomic or molecular systems, ions in crystals or gas molecules Rotation of dipoles	Infra-red or heat waves	Rubens Paschen Gratings and residual rays
7×10^{-5}	to 4×10^{-5} cm	Loosely bound outer valence electrons.	Visible light	Prisms, gratings, interferometers
4×10^{-5}	to 1.6×10^{-5} cm	Outer electrons more tightly bound	Ultra-violet.	Schumann and Lyman, gratings
1.6×10^{-5}	to 1.2×10^{-6} cm	Inner electrons of light atoms, or electrons in stripped light atoms, or shells next to valence shells of heavy atoms	Extreme ultraviolet or soft x-rays	Millikan and Bowen. Vacuum spectrograph, gratings.
1.2×10^{-6}	to 1.6×10^{-8} cm	Interior electrons of elements.	Very readily observed by all methods.	Thibaud reflected from glass gratings at grazing incidence.
1.6×10^{-8}	to 1.25×10^{-9} cm	Innermost electrons of atoms, shorter waves apply to heaviest atoms.	X-rays hard and soft, K and L series.	Crystal gratings, Laue, Bragg, also from $Ve = \frac{1}{2}mv^2 = h\nu$, $\nu = c/\lambda$.
1.25×10^{-9}	to 5.56×10^{-11} cm.	Nuclear electrons latter are hardest ones from radium C.	γ-rays.	Robinson, de Broglie, Ellis, Meitner from $\frac{1}{2}mv^2 = h\nu$, v from magnetic field

CHAPTER XXVII

THE PHOTOELECTRIC AND THERMIONIC EFFECTS

In the rapid adaptations of modern scientific discovery to industrial and practical purposes there are two more or less recently discovered and developed sets of electrical phenomena which are of wide application and of considerable importance. One of these is the so-called photoelectric effect and the other one is the thermionic effect. In a text of the scope of this one the widespread use of these phenomena demand their inclusion. However in keeping with the spirit of this text the subject will not be dealt with in its practical applications to any great extent, nor is there space to discuss anything but the underlying foundations of the many different applications.

PART I. THE PHOTOELECTRIC EFFECT

135. EARLY DISCOVERIES

In 1887 Hertz, the discoverer of the electromagnetic waves, noticed that the spark in his receiving set passed less readily (i.e., required a shorter spark gap) if the spark were screened from the spark in his primary sending set causing the oscillations. In a masterful research investigation, which could be chosen as a model of scientific analysis, he determined the nature of the phenomenon. He showed that ultraviolet light from the primary spark caused the spark in the receiving set to pass more easily if it fell on *one* of the electrodes, namely, the cathode. The radiations effective traveled in straight lines, were reflected, refracted and absorbed by certain substances, while others transmitted them. The wave lengths of light effective were well beyond the visible towards the violet end of the spectrum. He found that an electrical arc was more effective than the spark. The work was carried farther at once by Ebert and Hallwachs. The latter studied the action of the light from an arc on electrostatically charged bodies. It was found that the light falling on a negatively charged zinc plate connected to an electroscope caused the electroscope to discharge rapidly, while if the zinc plate was positive there was no discharge noticeable. The result was shown to be due to *ultraviolet*

light. Thus the *photoelectric effect* (literally light-electric effect) was discovered. The action of the light resides on the metal, the air being inactive, and is very much a function of the condition and nature of the metal on which it falls. A clean plate of an electropositive metal like aluminum or zinc was more effective than iron, and corrosion destroyed the effect entirely. Hallwachs finally concluded that negatively electrified particles travel away from the plate when it is illuminated, and that this occurs for an uncharged plate but is more effective if the plate is negatively charged. The residual charge on illuminating the uncharged plate was positive, proving the escape of

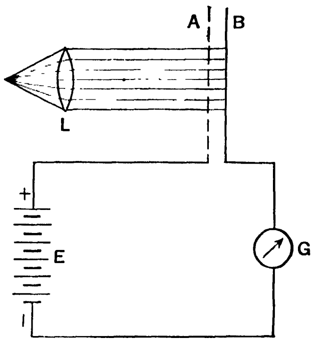


FIG. 131*a*.—Device for a Study of the Photoelectric Effect Using a Galvanometer.

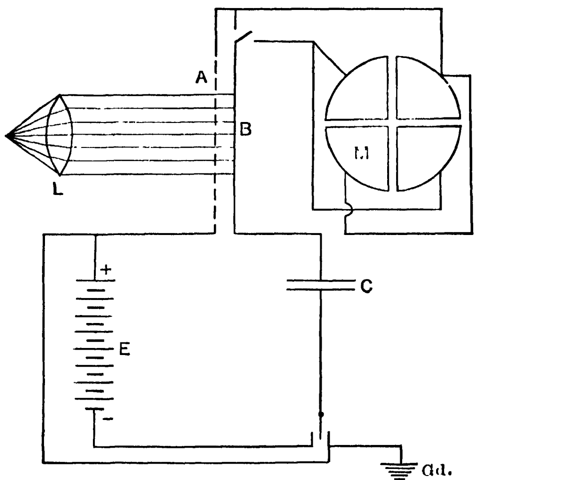


FIG. 131*b*.—Device for the Study of the Photoelectric Effect Using a Quadrant Electrometer.

negative particles to be the cause. For the study of this phenomenon Stoletow devised the simple circuits shown in Figs. 131*a* and 131*b*. In these the light from an arc or burning magnesium ribbon focused by a quartz lens L falls on a plate B after passing through a gauze A. A and B form two plates of a parallel plate condenser connected to the battery E through a galvanometer G, Fig. 131*a*, or an electrometer M, Fig. 131*b*. The plate B must always be negative. Fleming in 1909 found that if an alloy of Na and K was used on B while A was *Pt* an E.M.F. between A and B of 0.6 volt could be built up due to the flow of negative electricity from B to A, using scheme (*b*). This potential it was early discovered is the *contact difference of potential* (formerly called the Volta difference of potential) an intrinsic property

of the metals which was first observed by Volta in his studies on the galvanic cell. Fleming obtained a current of 5.4 microamperes, presumably with the battery and a galvanometer of resistance of 180 ohms. Righi had previously shown that the new photoelectric effect took place in vacuum with even better results than in the presence of gases. In 1890 Elster and Geitel showed that the effect in vacuum was reduced when a magnetic field parallel to the surfaces was used.

136. LIBERATION OF ELECTRONS BY LIGHT

The discovery of the electron by J. J. Thomson as a result of the investigations of the nature of the bluish streamers from the cathode causing x-rays in a highly evacuated tube led to a speedy identification of the negative electrical particles of Hallwachs with electrons. The experimental proof of this identity was made by Thomson in 1899. From then on the development of the subject owes much to P. Lenard in 1900 and thereafter. He checked Thomson's results, using the photoelectrons in vacuum by a slightly different method. The fact that the particles were electrons liberated by light led to the name photoelectrons given them today, and the effect is known as the photoelectric (or light-electric) effect. Lenard next studied the current as influenced by the potential between the illuminated plate or electrode and an opposing electrode in vacuum as a function of the potential difference between the electrodes. He found that at 0 potential difference he got a small but definite current. As he made the illuminated electrode more and more negative relative to the other electrode the current increased rapidly at first and then more slowly, reaching a maximum or " saturation " value at 100 volts or more, thereafter remaining constant. This indicated that as the electrical field increased more and more electrons that had been liberated by the light were drawn to the positive electrode, thus increasing the current, and that in his apparatus at 100 volts all electrons liberated were drawn out and contributed to the *saturation* value of the current. Together with other results this effect of the field indicated that the electrons were liberated in all directions and at different velocities. The reason that at low fields the currents were smaller lies in the fact that owing to this directional effect the weak fields let many electrons get back to the illuminated plate, or escape to the walls of the vessel. At higher fields the electrons were drawn more and more nearly directly across from cathode to anode by the field.

Lenard also found that as the illuminated plate was made slightly positive the current was reduced but did not stop until the plate was

positive by some 2.1 volts relative to the collecting plate. He correctly interpreted this as due to the fact that *some electrons were liberated at high initial velocities.* That at 2.1 volts no more electrons crossed meant that the fastest electrons which were headed directly for the collecting electrode from the illuminated electrode had only just enough energy to reach the collecting electrode against the field of 2.1 volts that retarded them. Since the electron carries a charge of e units ($e = 4.77 \times 10^{-10}$ E.S.U.) the work done on the electron in going from one electrode to the other against a field of V volts is Ve ergs of work (V in volts being properly expressed in E.S.U.). To be able to move against a field of V volts an electron of charge e must have a velocity v given by the relation $Ve = \frac{1}{2}mv^2$ where m is the mass of the electron in grams. Hence v in cm/sec $= \sqrt{\frac{2Ve}{m}}$. If V is 1 volt $= 1/300$ E.S.U., $e = 4.77 \times 10^{-10}$ E.S.U., and $m = 8.99 \times 10^{-28}$ grams, it is seen that an electron able to move up against a potential of 1 volt (or an electron that has been freely accelerated by a potential difference of 1 volt) will have a velocity of 5.948×10^7 cm/sec. In the modern physical terminology it is often more convenient to speak of electron velocities in V volts; that is, to describe the velocity as that acquired by an electron in a free fall across a potential of V volts. It is seen that the velocity is then actually given by $\sqrt{V}\, 5.948 \times 10^7$ cm/sec, V being expressed in volts. The velocity of the fastest of the photoelectrons observed by Lenard therefore must have been about 0.86×10^8 cm/sec. At lower positive potentials of the emitting plate slower electrons can cross to the collector as well as electrons whose directions are not normal to the collecting electrode, but whose components of velocity in that direction are such that they exceed $5.948 \times 10^7 \sqrt{V}$ cm/sec. Lenard and others before him also found that the presence of gases reduced the currents under otherwise identical conditions. This is due to a number of causes, namely:

(*a*) Change in the nature of the metal surfaces from which electrons were emitted as a result of adsorption of gas or chemical reactions.

(*b*) Reduction of energy of electrons by collisions with gas molecules and an angular scattering of electrons by the latter resulting in their ultimate return to the illuminated plate.

(*c*) By loss of electrons to molecules in the formation of slowly moving negative ions by electron attachment.

If actions *a* and *c* are absent (i.e., as in the case of argon or helium) the effect of gases should be merely to raise the saturation voltage. This does not seem to be the case and there is today no explanation of the failure.

It should be noted that the maximum velocities of the electrons observed depended not only on the *nature of the material of the illuminated electrode* but also on some condition inherent on *the combination of the metals* used in both electrodes if they were in any way different. This phenomenon involving both electrodes had been earlier observed and correctly attributed to *contact potentials* between metals, which act to accelerate or retard the photoelectrons and whose value must be determined and added to or subtracted from the applied potentials in any quantitative study. The contact potential is the potential difference which acting on unit charge represents the work required to move an electron from the one metal to the other through their surfaces against the surface electrical fields existing in the metals. These potentials are related to the values of C, μ and χ, discussed later on page 335.

In the discussion of all that has gone before it was assumed that the intensity and the nature of the light source was kept constant. It is not surprising that Lenard found that the saturation current varied with the light intensity, and that in fact the current was directly proportional to the light intensity. This law has been amply verified since and has been found to hold for light intensities which in total vary in the ratio of 1 to 10^5, although in any one experiment the study has not been made for a single range of intensities of more than 10 fold. It is, however, unquestionable that in the range studied quite accurate proportionality exists. It is, therefore, not surprising that in the great technical applications of today the proportionality between the electrical current in the photo electric effect and the light intensity should give a valuable means for converting varying light intensities to proportionately varying electrical intensities. It is in fact on this basis that the modern " talkies " as well as the new television and many other devices operate.

137. THE EINSTEIN PHOTOELECTRIC LAW

While it was early known that the frequency of the light played an important rôle in the photo effect the exact nature of the process was little understood. It was known that the effect was produced by ultraviolet light. The earlier workers Stoletow and Elster and Geitel found further that there was a limit on the red side of the spectrum below which ultraviolet light ceased to be active. It was further found that the wave length of the ultraviolet light which marked this threshold varied with the metal used. In other words there appeared to be a *long wave length* (low frequency) *limit* to ultraviolet activity

which varied with different metals used. It was also found that this threshold depended somewhat on the gas and the condition of the surface. Another early observation was the discovery of the effect known as photoelectric fatigue. It was observed particularly in the presence of gases, or residual gases, that the photoelectric effect varied with the time of illumination and that with prolonged illumination the quantity of current decreased. The fatigue is today little understood, and has been much reduced by the elimination of reactive gases and features causing the formation of surface films.

In 1905 a new era dawned in the investigation of the photoelectric effect. This was in part due to some observations of Lenard which indicated that the maximum velocity of the photoelectrons, as measured by the potential against which they could reach the collecting electrode, depended on the frequency of the ultraviolet light. This fact stimulated Einstein to discuss the effect on the basis of the then rather newly developed quantum theory. As a result of his attempt to find an accurate equation for black body radiation Max Planck in 1900 had arrived at the conclusion that a satisfactory equation could only be obtained by assuming that radiant energy (light) was emitted or absorbed in units or quanta. He had found that the size of this unit or quantum varied with the frequency ν of the radiation and was given by a whole multiple of the quantity $h\nu$, where h was a new universal constant of nature known as the unit of quantum action. As shown on page 69, ν is a frequency and has the dimensions of the reciprocal of time or $\frac{1}{T}$, while $h\nu$ has the dimensions of energy. Hence h has the dimensions of an energy times time, or ML^2T^{-1}. It is also of importance to note that mvr, the *moment of momentum*, has the dimensions of ML^2T^{-1} so that h has at the same time the dimensions and perhaps the properties of a moment of momentum, which is called action, as well as energy multiplied by time. Hence h is the unit of action. This peculiar dualistic property of the quantity h was destined to play an enormously important rôle in the hands of Bohr when he came to unraveling the spectra of the elements in relation to the new Rutherford theory of atomic structure. Basing his views on Lenard's observations concerning the frequency of the incident light and velocity of photoelectrons, and on a rather extreme interpretation of the notion of the rôle of the quantum in light, Einstein derived a mathematical equation expressing a relation between the frequency of the light and the electron velocity in photoelectric emission. This equation may be written

$$\tfrac{1}{2}mv^2 = eV = h\nu - P$$

In this equation m is the mass of the electron, v its velocity, e the charge, V the potential necessary to stop the electron (i.e., the volt equivalent of its velocity) and P the energy necessary to remove an electron from the metal with 0 velocity. That is, P represents the energy to just set an electron free from the interior and bring it through the surface of the metal. By bringing the frequency ν down to lower and lower values it is seen that $v = 0$ at $h\nu_0 = P$, where ν_0 is the frequency for which electrons emerge with no velocity. This frequency it turns out is the *limiting frequency* spoken of above, below which no more electrons can emerge; that is, ν_0 is the *red threshold* of the photoelectric effect. Einstein's equation then can also be written as $\frac{1}{2}mv^2 = h\nu - h\nu_0$. The quantum theoretical view point on which Einstein derived the equation was an extreme one, and he not only assumed absorption or emission in quanta as had Planck, but that the light was actually quantized in space. Later developments led him to conclude that his initial assumptions were untenable and hence that the deduction was not valid. With a peculiar irony of fate, however, the next researches of Ladenburg, Richardson and Compton and Hughes indicated that the energy of the photoelectrons was proportional to the frequency of the incident light, although the proof was not extensive enough to be conclusive. Finally in 1914–1915 Millikan and his pupils developed a technique for making a careful test of the law and *accurately verified it.* One of the great drawbacks in investigation had up to this time lain in the difficulty of getting *clean outgassed surfaces.* Using high vacua made possible by improvements in vacuum technique, and a new method of *shaving* the surfaces of his photoelectric metal in vacua, thereby getting fresh surfaces (the soft alkali and alkaline earth metals were largely used) Millikan was able to make a real test of the law. These experiments showed that the equation held very accurately. The source of error due to the contact potential was eliminated by proper control methods (i.e., independent evaluation of the contact potentials), so that the true energy of the photoelectrons could be measured, the values of ν_0 for the different metals were determined, and finally a new and for that period very accurate value of h was obtained. Within a year or two the work of Duane, and of D. L. Webster, established the law accurately for the photoelectrons from x-rays. The final and most accurate verification in the latter field came from the work of Maurice de Broglie and of Ellis between 1921 and 1925. In fact the equation is now so well established that it furnishes a method of obtaining the wave length and frequency of electromagnetic waves too short to be analyzed by crystal structure methods, such as, for instance, the gamma rays of radium.

138. THE INTENSITY OF PHOTOELECTRIC EMISSION

Having now the general principles of the photoelectric emission before us perhaps a few words might be said concerning some of the more detailed aspects. First mention might be made of the proportionality of photoelectrical current intensity and the light intensity. This law of proportionality holds up to current densities of the order of 3×10^{-7} amperes/cm^2. Further increase in intensity could cause apparent deviations unless care is taken to increase the field between source and collector, for with high current densities *space charges due to accumulations of large numbers of electrons can increase the saturation voltage necessary*, as we shall see later in connection with thermionics. Hence one must emphasize the necessity of the use of *saturation currents in all measurements.* For very small currents proportionality still holds down to 10^{-9} ergs/cm^2 incident light intensity of blue light, or 200 quanta per cm^2. Below this the currents must be amplified by ionization by collision, and emission is so slow as to be sporadic. The *average* emission even down to 10^{-11} erg/cm^2 per sec incident light intensity (2 quanta/cm^2 per sec) still is proportional to the intensity if measured so as to get an average value. The first electrons are liberated at the instant of illumination even over time intervals of 10^{-9} seconds. This fact caused some difficulty in the understanding of the action of light, for with the minute energies used it was impossible to understand how one electron in a metal surface could instantly (10^{-9} seconds) gain a whole quantum unless light went in particles. The wave mechanics has, however, removed this difficulty entirely.

When the study of the law of emission as a function of intensity is carried out for the use of light incident at angles differing from normal, care must be taken to remember that it is the *absorbed* light which causes emission. Hence as the proportion of absorbed to reflected light varies with the angle of incidence and reflection, we should expect a variation of the photoelectric intensity as a function of the angle of incidence for a constant light source. The work of Elster and Geitel and of Pohl showed that the change of emission with angle was quite accurately proportional to the *absorbed* light intensity for different angles.

It being understood that below a certain frequency ν_0 no light is able to call forth photoelectric currents; it is of interest to see how the current changes with frequency above the threshold. It is best in this case to plot current per unit light energy *absorbed* at various frequencies of incident light. The type of curve obtained for different

metals in such measurements is shown in Fig. 132. The curves begin at the long wave length limit, are concave upward, increasing more rapidly with ν than proportional to ν and approach a linear relation. It might be expected from classical reasoning that the curves should flatten out at the higher frequencies and show a portion concave to the origin. This variation has not been observed thus far experimentally *except with the alkali metals*, as will be seen later. It is also to be expected on the newer concepts of the electrical behavior of metals. It is, however, probable that the point of inflection in most measurements lies well above frequencies that can be controlled in the laboratory in such experiments. Various attempts have been made at deducing proper equations for this phenomenon, but until the recent work of Pauli and Sommerfeld on the theory of the nature of the electronic state in metals no success had been achieved. More recently on the basis of the Sommerfeld theory of the state of the electrons in metals Wentzel has made what appears to be a promising approach. The nature of the electron theory of metals is well beyond the scope of this text and hence cannot be included, although a word might be said concerning the nature of the new theory.

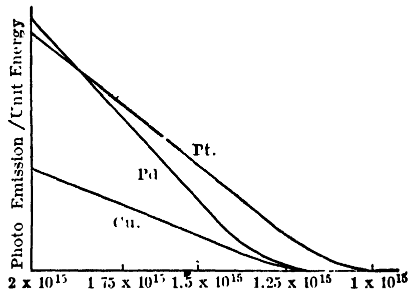


FIG. 132.—Photoelectric Current per Unit Incident Light Energy as a Function of Frequency of Incident Light in Cycles per Second.

139. THE SOMMERFELD-PAULI CONCEPT OF THE FREE ELECTRONS IN METALS

As a result of the high conductivity of metals and the optical properties of metals as stated in Chapter XIII the electrons in metals were assumed to be free in the crystal lattice of the metal in considerable numbers and to share in the thermal energy of the metal. In order that a *free* electron inside a metal which possesses 0 energy can escape from the metal a certain energy, of value C, must be given to it. That is, C is the *work* required to pull an electron out of a metal surface into free space against the attractive forces. It can be called the *true work function*. This work C is dependent on the nature of the metal (i.e., its atomic species and crystalline state). In general it represents a rather high energy value of the order of magnitude of that required to move an electron against a field of from 10 to 20 volts.

Were the electrons at *rest* in the metal and also *free*, then the photoelectric threshold would be represented by the relation that $C = h\nu_0$. Now in general electrons have long been assumed to be free in metals as a result of optical studies on the reflecting power of metals, as well as from considerations dealing with the peculiar constancy of the ratios of electrical and heat conductivities in metals known as the Wiedemann-Franz law. The estimated number of free electrons was of the order of one or more free electrons per atom present (i.e., in silver about 6×10^{22} electrons per cm^3).

If such electrons are *free* it was believed that they should be moving about among the atoms of the metal crystal lattice with an average energy of motion given by $\frac{1}{2}mv^2 = \frac{3}{2}kT$ (where T is the absolute temperature and k is the gas constant R divided by the number of molecules in a gram molecule), due to interchange of heat energies with the atoms of the lattice. At 0° C. this energy of motion amounts on the average to about 0.0356 electron volt of energy. Hence the work to remove an electron at 0° C. would not be C, but should on the average be $C - 0.0356$ volt. However, if the 6×10^{22} electrons in a cm^3 of Ag, for example, have this freedom of energy exchange with the atoms, they might be expected to take on a large amount of heat as the temperature of the body is increased, for each electron acts like an atom in energy exchange and there are 6×10^{22} electrons/cm^3. If the electrons acted according to classical theory this would increase the *atomic heats* of all metals by 50 per cent above the value *observed* by DuLong and Petit (atomic heats of pure crystalline metals = 6 calories). The fact that metals have the atomic heats given by DuLong and Petit's law thus indicates that *somehow the free electrons do not behave in a strictly classical manner.* It is at this point that Sommerfeld improved the situation by applying the *non-classical quantum statistics* of Fermi and Dirac, together with the famous *exclusion* principle of Pauli, to electrons in metals.

While we cannot here enter into the nature of the reasoning by which the law of distribution was derived the nature of the reasoning may be indicated. Normally, particles in an atomsphere (i.e., atoms in a gas, or electrons in an electron atmosphere) which are in thermal equilibrium have the velocities of their constituent particles distributed around a mean value represented by the famous Maxwellian distribution law. This says that out of a total of N particles in the gas the N_{dc} particles having a velocity between c and $c + dc$ is given by the law that

$$N_{dc} = \frac{4N}{\alpha^3\sqrt{\pi}}c^2e^{-\frac{c^2}{\alpha^2}}dc.$$

This law is pictured in Fig. 133 where $\frac{N_{dc}}{N}$ is plotted against c for values of $\frac{dc}{\alpha}$ equal to 0.1. In the figure α is the most probable velocity and is represented by the peak of the curve. This velocity α is determined by the absolute temperature T of the gas, and is related to T by the relation

$$\frac{m}{2}\alpha^2 = kT = \frac{R_A}{N_A}T.$$

Here m is the mass of the particle (atom or electron) and k is the gas constant R_A per mole divided by the number of atoms in a mole.

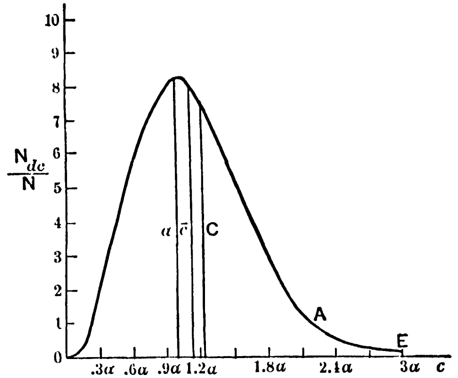


FIG. 133.—Maxwell's Law for the Distribution of Velocities among Molecules. Ordinates Are $\frac{N_{dc}}{N}$ in Per Cent, Abscissae c for $\frac{dc}{\alpha} = 0.1$.

It is for some purposes more convenient to divide the velocity c into components u, v and w along the X, Y and Z axes, and to express the chance that a particle has simultaneously velocity components between u and $u + du$, v and $v + dv$ and w and $w + dw$ as

$$f(u, v, w)\,du\,dv\,dw = \frac{du\,dv\,dw}{\alpha^3(\pi)^{3/2}}\,e^{\frac{-\frac{m}{2}(u^2+v^2+w^2)}{kT}} = \frac{du\,dv\,dw}{\alpha^3(\pi)^{3/2}}\,e^{-\frac{u^2+v^2+w^2}{\alpha^2}}.$$

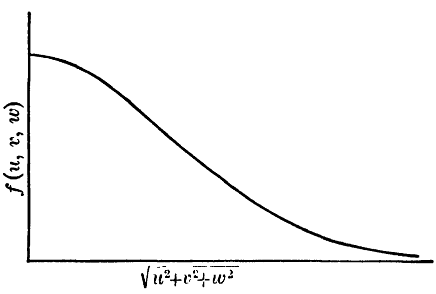


FIG. 134.—Chance that a Molecule Has Simultaneously the Velocity Components u, v, and w Plotted as Ordinates against the Combined Velocity $\sqrt{u^2 + v^2 + w^2}$ as Abscissae.

This form of the function is shown plotted against $c = \sqrt{u^2 + v^2 + w^2}$ in Fig. 134.

This law holds very well for gases, particularly as gases do not exist as such at very low temperatures. Under the conditions which might obtain at temperatures near absolute zero in gases and which do obtain for electrons in metals at room temperatures this law cannot be realized. The reason for this lies in the fact that it appears that

even the translational energies of a gas or an electron assemblage can be *quantized* by using the principle of action; for the conditions in the gas can be accurately pictured in terms of the momenta mu, mv and mw along the three axes and by the positions x, y, z of the separate particles, and the proper combinations of momenta and distances give action which can be quantized. The products of the momentum components along x, y and z and the space coordinates which lead to the action units that are quantized can be assigned different values depending on how many quantum units they contain, so that a picture of a gas at any temperature can be derived by counting up the various combinations of quantized action states which can go to describe the state of the gas, as regards the momenta and positions of its component parts. The statistics of Fermi and Dirac permit one to do this for the average state of an ideal gas, and add a restriction that *only* 1 *atom, or* 2 *electrons* (as there are two kinds of electrons with right and left handed spin), *may at one time occupy each one of the possible quantum states describing the gas.* The *number of quantum states in such a picture is limited and finite.* If now the *concentration of atoms or electrons is high,* space is so limited per atom or electron that there would be at the low temperatures *more than* 2 *electrons for each state, if we assumed the ordinary Maxwellian distribution* given above; for all the electrons or atoms are crowded into the narrow limits under the bell-shaped curve. But since there can be only 2 electrons in each separate state the Fermi-Dirac statistics give us a new distribution law which must vary in form with the density of the electron gas, and with the temperature in a fashion to be described below. The new law may be written as follows:

$$f(u, v, w)\, du\, dv\, dw = \frac{2m^3}{h^3} \frac{du\, dv\, dw}{e^{\left[\frac{m}{2}(u^2 + v^2 + w^2) - \mu\right]/kT} + 1}.$$

In this equation all symbols are as before except that Planck's h appears here as a result of quantization, and the 2 refers to the fact that 2 electrons may occupy each cell. The quantity $\mu = \left(\frac{3n}{\pi}\right)^{3/2} \frac{h^2}{m}$ to a first approximation is the important new factor which appears. In this term n represents the number of electrons per cm^3 of the metal. For values of $\frac{m}{2}(u^2 + v^2 + w^2)$ (the kinetic energy of an electron) less than the critical constant μ, it is seen that for relatively low values of kT (i.e., lower temperatures T) the exponent of e has a

negative sign and is large. Hence the exponential term vanishes in comparison to 1 and we have a distribution law of the form

$$f(u, v, w)\, du\, dv\, dw = \frac{2m^3}{h^3}\, du\, dv\, dw,$$

which is constant. At the value of a velocity

$$v_0 = \sqrt{\frac{2\mu}{m}} = \sqrt{u^2 + v^2 + w^2}$$

corresponding to $\mu = \frac{m}{2}(u^2 + v^2 + w^2)$, for small values of kT, the exponential term becomes 0 and the exponent changes sign rapidly, increasing as v rises above v_0. For higher values of v the 1 in the equation may be neglected, and $f(u, v, w)\, du\, dv\, dw$ drops sharply in an exponential manner to 0, remaining at 0, following however a curve of the Maxwellian form shown in Fig. 134. If T is very nearly 0 the new distribution curve is that shown in Fig. 135 by the full rectangular lines, the drop occurring at the velocity v_0 above. The higher the value of μ (i.e., the greater the electron density), the higher the value of T for which a given rate of fall occurs, and the higher lies the critical velocity v_0, or the critical energy $\frac{m}{2}v_0^2$. The interpretation of this curve now becomes obvious, for it is to be remembered that at very low temperatures (for T near absolute 0) there were so many low energy states among the electrons that they exceeded the number allowed according to the arrangement of states given by the Fermi-Dirac statistics. These electrons still exist in a free state but are bound in an energy distribution which gives them far more energy than the temperature of the metal permits, because they cannot occupy lower energy states and hence must go to the higher ones that are vacant. Accordingly, in a metal at low temperatures while there are a few electrons in equilibrium with the temperature which can exchange energy according to the usual laws, there are a great many electrons occupying positions of high energy content which are unable

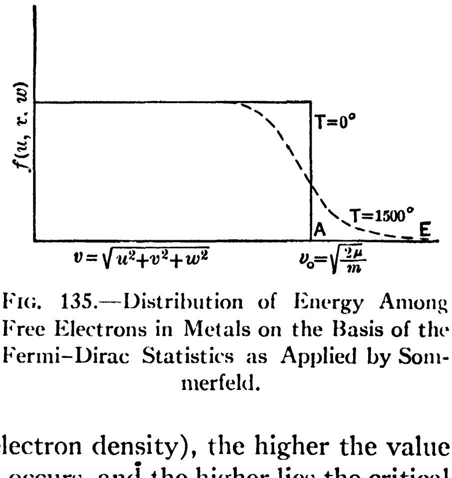


FIG. 135.—Distribution of Energy Among Free Electrons in Metals on the Basis of the Fermi-Dirac Statistics as Applied by Sommerfeld.

to share their energy with the atoms of the metal. The electrons in this peculiar state are called *degenerate.* At absolute zero all the electrons are degenerate, and the distribution of energy among the degenerate electrons is that given by the rectangle of Fig. 135. In this condition the number of electrons of velocity components du, dv, dw, is constant up to the critical combined velocity v_0,

$$v_0 = \sqrt{(u^2 + v^2 + w^2)_0},$$

and then 0 thereafter.

As T is increased the exponential function for velocities less than v_0 begins to be appreciable, and the exponential term with negative exponent adds to the 1 and makes the curve fall before v_0 is reached. For values of the velocity $\sqrt{u^2 + v^2 + w^2}$ greater than v_0 the exponential increases more slowly the greater kT, and hence the fall is more pronouncedly asymptotic and for a metal with 1 electron per atom at 1500° absolute the end of the distribution law curve takes on the appearance shown by the curved dotted line of Fig. 135. At a $\sqrt{u^2 + v^2 + w^2}$ with an energy greater than μ for large kT the exponential term is greater than the 1 and the distribution law approaches a law of the form

$$f(u, v, w)\, du\, dv\, dw = \frac{2m^3}{h^3} e^{\frac{-\frac{m}{2}(u^2+v^2+w^2)}{kT}} du\, dv\, dw,$$

which is closely similar to the form of the Maxwellian law as given in Fig. 134. In other words, the new law at high temperatures even for metals takes on the Maxwellian form.

We can thus summarize the situation by saying that at ordinary temperatures in a metal the high electron density leads to a distribution in which most of the electrons are free but degenerate, having a high energy whose limiting value is μ. As temperatures increase, more and more electrons become free to engage in energy exchanges, and the distribution function becomes modified in form, the higher velocity limit conforming more and more to the Maxwellian distribution.

140. THE INFLUENCE OF THE SOMMERFELD-PAULI THEORY ON PHOTOELECTRIC EMISSION

It is now of interest to inquire into the magnitudes of these quantities and to see how this new limiting energy will alter the photoelectric effect. For a metal like Ag with 1 electron per atom n takes the value of 6×10^{22} electrons per cm^3. With this value of n one can

calculate a value of μ which gives v_0 as about 5.5 equivalent electron volts of energy. Hence while at room temperature the average thermal energy of electrons is about .0356 electron volt there are present in the metal an appreciable number of electrons with an energy of 5.5 electron volts. Now while these electrons cannot partake of thermal equilibrium relations, they are *free*, and can respond to the vibrations of ultraviolet light. Hence if ultraviolet light of frequency ν_0 strike a metal with an absolute work function C, but with degenerate electrons of maximum energy not .0356 volt but equal to μ volts, the electrons will be emitted with zero energy if $h\nu_0 = \chi = C - \mu$. Hence χ gives us the actual *observed value of the work to pull an average* electron out of the metal. If the problem of the current as a function of the frequency ν be handled from the standpoint of the above suppositions using the wave mechanics, Wentzel shows that for $h\nu$ greater than C (i.e., all electrons can be removed by the light) the current i is proportional to $\nu^{-7/2}$. For values of $h\nu$ between C and χ the current i is proportional to

$$\nu^{-7/2}\left[1 - \left(\frac{C/h - \nu}{\mu/h}\right)^{3/2}\right]$$

while at a value of $h\nu$ below χ (i.e., below ν_0) there is no current. In the region between C and χ at low temperatures it is the degenerate electrons that give rise to i. Hence since most measurements are made in this region it is not surprising that *the photoelectric current is practically independent of temperature*, for most electrons are degenerate up to temperatures of 1000° C. or more. At temperatures where the electrons cease being degnerate there should be a definite effect of temperature. The effect is masked, however, by the thermionic effect in this region.

From the second equation of Wentzel, which in general covers the usual conditions, it is seen that the current reaches a maximum value and then decreases. It now happens that this maximum lies in the far ultraviolet for all except one group of metals, that is, for the alkali group. It is possible to calculate μ for data taken from the optical constants. χ is observed experimentally so that both C and μ are available. From these data one can calculate the equivalent wave lengths λ in Ångström units for the constants C/h, μ/h and ν_0 for two metals. They are

	$\lambda_{(C/h)}$	$\lambda_{(\mu/h)}$	λ_0
Na..........	2490	3930	6800
Cu..........	1110	1760	2660

From this table it seen that for Na the limiting wave length λ_0 is in the red of the visible spectrum, and the frequency corresponding to the energy C, $\lambda_{(C/h)}$, lies well in the ultraviolet, but in a region easily achieved. For copper the threshold wave length λ_0 is already in the ultraviolet, not far removed from the strong ultraviolet mercury line 2537, while C lies in the experimentally difficultly accessible *extreme ultraviolet.*

It is not strange then that the early work with the alkali metals showed a distinct maximum for the photoelectric emission. Pohl and Pringsheim further showed that this selective photoelectric effect, as it was called, depended on the direction of vibration of the light. The effect is a maximum when the electrical vector of polarized light is parallel to the plane of incidence (i.e., when the light vibrations have marked components perpendicular to the surface). The metals show peaks of activity at approximately the following wave lengths, Rb at 5000 Ångström units, K at 4500, Na at 3400, and Li at 2800. The theory of Wentzel predicts just such an action with a dependence on the electrical vector as indicated, though at present in approximate form only. An idea of the magnitudes of the quantities λ_0, μ expressed in equivalent volts of energy and as an equivalent wave length by the relation $3 \times 10^{10}\, h/\mu$, χ in equivalent volts observed from the photo

Substance	Long wave length threshold in Å units	μ expressed in equivalent wave length in Å u. by $3 \times 10^{10}\, h/\mu$	μ in equivalent volts.	χ observed in Volts		Approximate value of C in equivalent volts
				From photo effect	From thermionic effect	
Li.	5200–5160	2630	4.7			
C	2565–2615			4.72–4.81	4.3	
Na.	5830	3930	3.14	1.80–2.12	1.8	5.0
K.	6120–7200	6050	2.04	1.2 –2.02	0.46–1.55	4.0
Cu	2665–2750	1760	7.01	4.07–4.63	3.85–4.00	11.5
Ag.	3210–3390	2250	5.5	3.75–4.05	2.60	9.5
Rb.	10000	6980	1.77	1.2	1.45	2.6
Cs	10000	8150	1.52	1.2	0.7 –1.36	2.2
Zn.	3020–4010			3.08–4.10	3.02	
Au.	2600–2730			4.33–4.75		
Fe.	2870–3150			3.92–4.36	4.04	
Hg.	260–304			4.05–4.75		
Pt.	1850–3020 best 2840			3.63–6.5	5.0 –6.0	
W.	2300–2735			4.52–5.36	4.31–4.53	

effect and the thermionic effect, and finally $C = \mu + \chi$, is given in the accompanying table. It is to be noted among other things that the quantity χ comes both out of photoelectric and thermionic effects, a result to be expected from this theory.

141. FURTHER FACTS CONCERNING PHOTOELECTRIC EMISSION

The distribution of the velocities of photoelectrons liberated by different wave lengths from different metals is of the same general form for all metals and frequencies, and differs only in the magnitude of the currents found, according to Ramsauer. The form of the curve is a sort of bell-shaped curve when energy of the emitted electrons is plotted against the number of electrons emitted. The peak, or most probable energy in these curves, is between $\frac{1}{3}$ and $\frac{1}{2}$ the maximum emergent energy. Such a curve of velocities of *emitted* electrons might indicate the form of the energy distribution of the degenerate electrons in the metal, but is however involved with the question as to the initial direction of emergence, and the nature of the surface layers. The results of Ramsauer do not appear to be of very general application for substances other than certain pure metals, and surface films play a very disturbing rôle.

The photoelectric effect occurs not only for metals but for a large number of non-metallic substances. In these cases the light does not only have to liberate the electrons from the solid or liquid surface, but is also required to first set free electrons in the substances, for non-metallic substances are characterized by the absence of free electrons in any quantities. In general very short wave lengths are required in such cases, and we see the long wave length thresholds required for substances like NaCl, KCl, CuCl, etc., lying in the region of 1800–2200 Ångström units. The action of light on AgI, on S, and on anthracene indicates internal liberation of electrons in the visible region with external emission only in the ultraviolet. In all cases the currents are small, as relatively little light is absorbed near the surface (in comparison to metals), so that many electrons cannot get out. The vapors of the metals are also photoelectric, and it has been shown by Williamson, Mohler, Lawrence, and others that ultraviolet light liberates electrons in vapors of the alkali metals and Hg. In gases like N_2, O_2 and H_2 the effect should also be observed, but the long wave length limit lies so far in the ultraviolet as to preclude detection. The effects are also small owing to the great transparency of gases and vapors to light.

Finally it might be stated that in recent years the action of a high

external electrical field on the photoelectric effect has yielded the fact that the threshold frequency is shifted to longer wave lengths if a powerful field is allowed to act to pull the electrons out, as shown by Lawrence and Linford. This is in agreement with the theory of Schottky as to the nature of the electrical fields at the surface of metals, and with Wentzel's theory of the metallic surface forces.

The photoelectric effect is used wherever it is desired to convert changes of light intensities to electrical or mechanical effects. It can with certain restrictions be used for *measuring light intensities*, provided *the selective actions of some wave lengths* are corrected for. It can be used for producing electrical effects proportional to varying light intensities. It is widely used for recording speech in talking moving pictures. Here the varying light intensities produced by the reception of sound waves on an electrical receiving system sensitive to sound vibrations are photographed on the same films recording a moving picture. This insures synchronism between speech and action. After developing the film the varying intensities of the light record on the film are projected onto a photoelectric cell by means of a beam of light passed through the film. The electrical response of the cell is reproduced as sound in a loud-speaker through electrical amplifying devices. Photocells are also used in the transmission of pictures of objects by radio in television. The light intensities at various points on the object being viewed are in rapid succession thrown onto a photocell which in turn modifies the intensities of signals emitted by a radio sending device. At the receiving end the signals are amplified up, synchronized with the sending device, and used to produce light intensity variations by changes of current in an appropriate type of neon lamp. These varying light intensities are thrown as spots of light in rapid succession having appropriate intensities onto a screen where the persistence of vision integrates the flashes into a picture. The photoelectric effect is used widely in the control of any phenomena, electrical or mechanical, by light signals.

The advantages for use of the photoelectric effect in light intensity measurements are:

(1) The close proportionality between light intensity and current.

(2) The very rapid time response of signals with no apparent inertial effects.

(3) The extreme sensitivity of the effect. (The effect with amplification is possibly more sensitive than the eye for light detection. The eye can detect light intensities of the order of one quantum of blue green light per mm^2 per second, an effect easily detectable with the photocell.)

The disadvantages of the photoelectric effect in light intensity measurements are:

(1) The photocells are all *selective* in their action. Certain wave lengths are far more effective than others. Hence great care in the adaptation of such cells and the interpretation of results obtained therewith is required.

(2) The unsatisfactory degree of constancy of emission. The photoelectric emission varies with time and requires constant verification.

(3) The unsuitability of the phenomenon for measurements of wave lengths greater than 6000 Ångström units.

The commercial photoelectric cells have either a central anode or a central cathode. Cells with a central anode are most frequently used. The cell consists of a glass or quartz (for ultraviolet light) bulb coated on the inside with a sensitive conducting layer of a photoelectric metal. This is usually one of the alkali metals, very often in the form of a hydride (i.e., combined with hydrogen). This coating is removed from a given area of the cell to make a window for the admission of the light. The central anode usually in ring form is insulated from the layer of metal, and both anode and metal coating or cathode are connected to outside contact points by sealed-in leads. The cells are usually highly evacuated, except for certain cells which have a residual gas for amplifying the currents by ionization by collision. The usual cells employ enough potential between cathode and anode to give a saturation current, i.e., about 100 or so volts. The gas-filled cells use a much higher potential (just below sparking) to cause the electrons emitted by the metal to multiply the current by ionization by collision. The alkali hydride cells are most sensitive in the visible but are not very constant. The hydride is formed by a glow discharge between cathode and anode in the presence of H_2 at a pressure of a mm or so of mercury. The gas is later pumped out. Constant cells of very high efficiency are given by Zn-Cd cells if ultraviolet light can be used. Straight alkali metal cells are also used for the visible end of the spectrum and, if well made, are fairly constant.

To measure light intensities the cell can be used as indicated in Fig. 136. S is the cell, C the cathode, A the anode. B is a battery of from 90 to 200 volts, and G is the galvanometer while R is a resistance of some 5000 ohms.

In Fig. 137 an amplifying circuit and photocell are shown for working a relay R in response to light signals. B is a battery of 90–135 volts, M_1 is a grid bias variable resistance of 10^4 ohms, with a battery b for giving a bias to the grid of the 3 electrode tube. The

cell is S, the cathode C being connected to the grid G of a 3 electrode amplifier tube. M_2 is a resistance of 1–10 megohms between the bias and the cathode C. The anode A is connected through a milliammeter A_1 through the relay coil R and to the plate P of the 3 electrode tube, the positive pole of the main battery B also going to A. B_1 is the filament lighting battery of the filament F of the 3 electrode tube which is attached to the negative pole of the battery B, and the positive pole of the bias b through the variable resistance M_1. On illumination the positive charge on C due to loss of electrons overcomes the negative bias on b which is set to interrupt the electron current

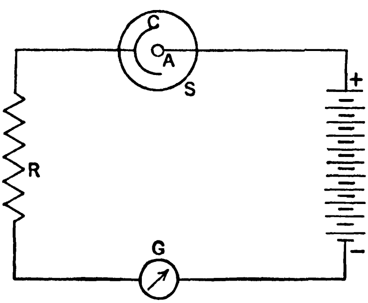


FIG. 136.—Simple Circuit for the Use of a Photoelectric Cell.

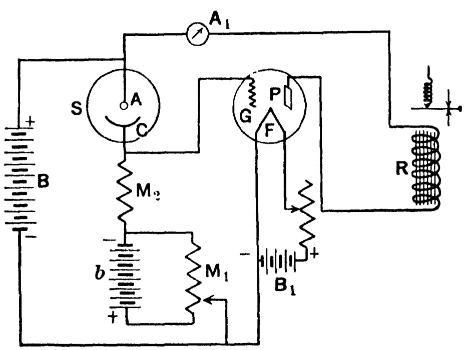


FIG. 137.—Amplifying Circuit Used to Work a Relay by Means of a Photo Cell.

from the filament F to the plate G. The instant the bias is neutralized the current flows in the 3 electrode tube and operates the relay.

PART II. THE THERMIONIC EFFECT

142. THE DISCOVERY OF THE THERMIONIC EFFECT

We can now turn to the second of the two phenomena to be discussed, namely, the thermionic effect. The early experiments on static electricity of Du Fay, 1737, Watson, Priestley, and others indicated that the air in the neighborhood of hot bodies has the power of conducting electricity. These early observations were not pushed any further and it was not until Becquerel in 1853 observed that air near objects at white heat was unable to insulate, even for a few volts potential difference, that any real advance in the study of this field was made. This fact was extended by Blondlot to include as low a potential difference as 0.001 of a volt. The latter also found that the current was not proportional to the potential difference for small

potentials, a very important fact as we shall later learn. It was Elster and Geitel who made the first systematic investigation of the phenomenon in 1880. It will be recalled that these same men were also pioneers in a study of the photoelectric effect. They heated wires by an electrical current and measured the current conducted to a neighboring electrode as a function of the potential between hot wire and electrode. It was observed that a current was established in the absence of any potential difference. This current built the potential between the electrode up to a certain value and then ceased, except in sufficient measure to maintain the potential built up. The effects obtained varied widely. With Pt in air the potential was positive at low temperatures (dull red heat), increased to a maximum at red heat, and fell to near zero at white heat. In vacuum the potential at higher temperatures not only went to 0 but became progressively more and more negative. Thus the wires appeared to give off positive electricity at low temperatures and negative electricity at high temperatures. The effects varied with the metals used. Cu gave mostly positive charges, carbon exclusively negative charges. Branly, approaching the problem with a different technique, confirmed the results of Elster and Geitel. Edison in his early work on the carbon filament lamp found that the filament became positive while an independent electrode in the lamp bulb became negative when the filament was sufficiently hot. This was analgous to the results of Elster and Geitel and was proven so by Preece and by Fleming in 1885 to 1896. J. J. Thomson's work leading to the discovery of the electron and the investigations on positive and negative gaseous ions led to the opinion that the very hot filaments gave rise to negative ions, especially in the presence of gases. It was then believed that the hot metal acting on the gas ionized the gas, and McClelland, a pupil of Thomson's, in 1899 showed that in fact the air drawn from the neighborhood of the hot filament did contain negative ions. This was proven by a measurement of the velocity of the ions, these ions being somewhat slower than the normal ions in air. This was due to a faulty technique, for now the ions formed in dry air by a hot filament are known to have the same velocity as normal negative ions. McClelland found that the current from a hot wire increased with the potential applied, finally reaching a *constant saturation* value. J. J. Thomson, as might be expected, measured the ratio $\frac{e}{m}$, of charge to mass for the ions from a hot carbon filament in 1899 and showed the *negative carriers to be electrons*, or the same as the constituents of cathode rays. Hence the negative currents from hot carbon, and from the hotter metals like

Pt, were electrons liberated in vacuum and thus not a product of gaseous disintegration produced by the hot wire as previously assumed. The ions observed by McClelland were simply electrons that had attached to air molecules or vaporized platinum particles. *The electrons are emitted directly from the incandescent bodies at high temperatures in vacuo.* By 1906 J. J. Thomson had shown by magnetic measurements, similar to those used in $\frac{e}{m}$ determinations, that the positive electricity emitted by Pt at lower temperatures was on particles of a molecular mass (either ions of Pt or lighter ions attributed to the gas present), and that some heavier ions were also present as observed by McClelland. These are now attributed to large positively charged solid particles of platinum torn off, or liberated, from the hot wire as atoms that condense to solid particles.

143. THE LAWS OF THERMIONIC EMISSION

By 1901 the pioneer work on thermionic emission which has made possible most of the modern advances in the field of thermionics, and which netted its author the Nobel prize, was begun by O. W. Richardson. This work was an experimental investigation of the variation of the electron emission from hot bodies as a function of temperature. The problem was a difficult one from a technical point of view, particularly at its inception. The body to be studied is if possible used in the form of a small wire W supported in the axis of a cylindrical tube by heavy leads L as shown in Fig. 138. The wire is surrounded by a cylinder C of length greater than the wire and coaxial with the glass tube. The cylinder and heavy leads are passed out through the glass tube by glass to metal seals, and a side tube A is provided for exhausting the air. Owing to intense heating by the filament, all occluded gases must be baked out and removed by heating the glass to near its softening point, heating the metal parts to incandescence by an induction furnace, and bombarding the metal parts by a heavy

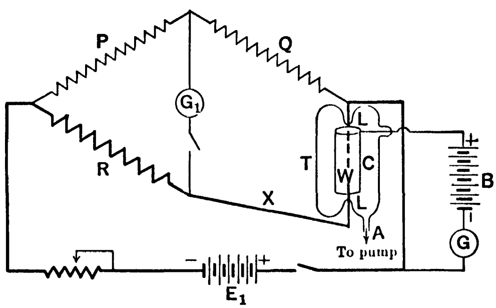


FIG. 138.—Arrangement for a Study of Thermionic Emission.

discharge of electrons. These operations are carried out while the tube is connected to a powerful high vacuum pump. Such a tube is then constant in further investigations and should yield reliable results. Modern tubes are sealed off from the pump on exhausting and are often kept gas free by distilling a small amount of reactive metal like an alkali or Mg into the tube to act as a "getter" for gases given off. The temperature of the incandescent filament is controlled and kept constant by measuring the filament resistance, changing the heating current as the temperature changes. The filament is made one arm of a Wheatstone bridge (X) of Fig. 138, the other resistance in that arm being R and having such a nature that it can carry the heating current without marked change in temperature. The other arm of the bridge has the ordinary ratio resistances P and Q set to give a balance with the hot wire at the proper temperature. The potential E_1 operating the bridge is the same one that gives the heating current. Any change in temperature of the filament changes the resistance and throws the bridge out of balance. The filament temperature can be determined by an optical pyrometer, or by a thermocouple. Calibrating the bridge to read temperature may also be accomplished by observing the melting points of specks of fusible substances placed on the filament, the bridge current being noted as each speck melts, so that the current can be plotted to indicate temperature. In this fashion by connecting the tube as indicated in Fig. 138, a study of emission as a function of temperature can be made.

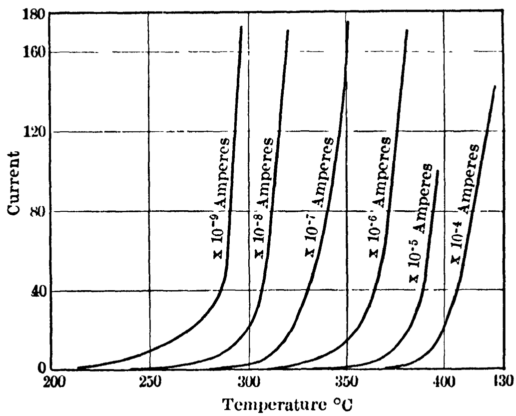


FIG. 139.—Richardson's Curves for the Thermionic Emission as a Function of Filament Temperature.

The battery B furnishes the potential driving the current through a galvanometer or electrometer G. The *potential* of B must be such as to give the *saturation current* from the filament, as it had been early observed by Blondlot that the current varied with potential for low values of the potential across the tube. It will be found that the potential to give this saturation increases with the temperature of the filament and hence as the current increases. A characteristic set of curves found by Richardson for Na from 217° to 427° C are shown in

Fig. 139.* It is seen that the saturation electron current from a hot filament rises exponentially with the temperature of the filament, following a law deduced by H. A. Wilson and Richardson theoretically as early as 1901, and later further developed by Richardson. The equation was developed on the basis of the electron theory of metallic conduction which evolved in the work of Drude, Riecke and Lorentz, following J. J. Thomson's discovery of the electron. As was stated in the early part of this chapter, the electrons were believed to be free in metals in considerable numbers and to be in thermal equilibrium with the atoms of the substance at a temperature corresponding to that of the body. If one assume that the energy of the electrons is distributed according to Maxwell's law, then as the temperature rises the velocities and energies of the electrons increase as described before. It will be noticed on observing the curve of Fig. 134, showing Maxwell's distribution for electrons emerging from a surface, that there are always present in the distribution some electrons of very high kinetic energy. If now the energy of motion of those electrons which are moving towards the surface of the metal, and near it, exceed the potential energy of the force field holding the electrons inside the metal, the electrons will, just as do atoms inside a liquid or a solid, escape through the surface (i.e., evaporate). Hence on heating a metal with free electrons to a high enough temperature so that the energy of a portion of the faster electrons (i.e., those in the tail AE of the distribution curve, Fig. 134) exceeds the so-called retarding potential at the surface, these electrons will escape, i.e., the hot metal will emit electrons. In close analogy with the kinetics and thermodynamics of vapor molecules escaping from a liquid surface (i.e., evaporating) Richardson and Wilson developed the equations for emission. There are two equations resulting from somewhat different assumptions as to conditions in the metal. One may designate by χ the energy required to remove an electron from the metal, i.e., the so-called *work function* mentioned before. For the case that the number of free electrons n per cm^3 is independent of the absolute temperature T, and that χ is independent of T, the equation becomes

$$i = n\epsilon\sqrt{\frac{k}{2\pi m}}\,T^{1/2}e^{-\chi/kT} = A_1 T^{1/2}e^{-\chi/kT}.$$

Here ϵ is the charge on the electron, m is its mass, and k is the Boltzmann constant $\dfrac{R_a}{N_a}$.

* The author wishes to acknowledge the courtesy of the editors of the Proceedings of the Royal Society of London, of Messrs. Longmans, Green and Co., and of Professor O. W. Richardson for permission to reproduce these curves.

If χ is independent of T, while n is proportional to $T^{1/2}$ the equation takes the form

$$i = A_2 T^2 e^{-\chi/kT}.$$

In more recent years Dushman has deduced the theory more rigorously and obtained the equation

$$i = \frac{2\pi k^2 m\epsilon}{h^3} T^2 e^{-\chi/kT}.$$

Here ϵ is the charge on the electron while e is the base of the natural system of logarithms.

The new statistics of Fermi and Dirac as recently applied to the problem of electron emission by Sommerfeld and Fowler gives the current i as

$$i = \frac{4\pi m\epsilon k^2}{h^3} DT^2 e^{-\frac{C - \mu}{kT}}$$

where D is the reflecting factor of the surface for electrons and $C - \mu = \chi$. This differs from Dushman's theoretical equation only by the factor 2 brought in by the electron spin (2 electrons for each energy state), and by the uncertain value of the coefficient D for internal reflection of electrons by the surface which, however, may cancel the 2 by taking on the value one-half. The quantity D enters the problem here as a result of the application of the newer wave mechanics and cannot be further considered. This equation, as the notation indicates, takes account of the high energy μ of the degenerate electrons.

It is seen then that in general except for the first equation the law of thermionic emission may be written

$$i = AT^2 e^{-\chi/kT} = AT^2 e^{-b/T}.$$

It may be recalled that χ is *not the real work function* but the *apparent* one, C being the real work function, and χ being of importance because the upper energy μ of the degenerate electrons reduces the work to get out of the metal: χ is given by $\chi = C - \mu$, and indicates the energy of thermal agitation beyond μ which is required by electrons in order to escape. In most work the constant χ/k is written as b or b_2. On the basis of Dushman's derivation the constant A has been evaluated as 60.2 amperes/cm^2 degree^2, a value agreeing within the limits of certainty with the constant of the Sommerfeld theory. As seen before $\chi = C - \mu$, and hence b is a characteristic of the metal used and depends on the work to pull electrons out of the metal. C is

the absolute work function and μ is the limiting energy of the fastest electrons in the Fermi distribution.

The equation has been tested numerous times, one of the latest and best tests having been in the work of Davisson and Germer, and later Germer. For a test of the equation over large ranges of i owing to the exponential term it is simpler to plot log i against $\frac{b}{T}$. If the equation is correct the resulting curve should be a straight line whose slope at once gives log AT^2. Such a curve, taken from Germer's paper in the *Physical Review*, Vol. 25, 805, 1925, is shown in Fig. 140. The way in which the observed points lie on a straight line indicates that the exponential law is accurately fulfilled even for a range of values of i from 1 to about 10^{10}.

The value of A obtained from the data for the element W is in good agreement with Dushman's value of A. It was early pointed out by Richardson, however, that in attempting to distinguish between the T^2 and the $T^{1/2}$ in the first two equations deduced, the rapid variation of i with $e^{-b/T}$ masked the effect of T^2 outside the exponential to such an extent that no decision between the two equations could be arrived at. The same condition holds for the value of A. This lies in the fact that between 1000° absolute and 2500° absolute i for W varies by a ratio of 10^{15} while $T^{1/2}$ varies by 1.58 and T^2 by only 6.25. What really emerges from theory and experiment is clearly that an equation of the form $i = AT^x e^{-b/T}$ fits the facts to a high degree of precision, with a good degree of probability that $x = 2$, and that the value of A observed by Davisson and Germer and computed by Dushman is of the right order of magnitude. The values of b observed are fairly accurate and consistent for similar surfaces and further agree within experimental limits with the values of $\chi/k = \frac{C - \mu}{k}$ determined by the photoelectric effect.

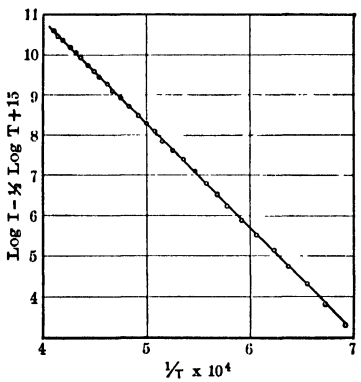


FIG. 140.—Germer's Test of Richardson's Equation for Thermionic Emission.

Hence we may be assured that the Dushman theory and the newer Sommerfeld theory are correct in general outline, though the more minute details are uncertain.

The relations established, however, have important implications, for they show that certain theoretical assumptions are confirmed. It was early shown by Richardson theoretically that if a Maxwellian distribution of velocities exists for electrons inside the metal they emerge through the surface with the Maxwellian distribution maintained, even though in escaping from the surface they have done work in escaping. This same condition holds for an evaporating liquid both the molecules of the liquid and the escaping vapor having the Maxwellian distribution of velocities, the average energy of the vapor in *equilibrium* being that of the liquid. Now the test of this expected behavior on the part of the electrons from a hot metal lies in the proof of the exponential law as predicted by the equations above. This is seen to have been verified to a high degree of precision, for the plot of log i against $\frac{b}{T}$ is accurately a straight line over a tremendous range. However, in the light of the newer Sommerfeld theory the electron energy distribution in the metal is not Maxwellian, but is given by the Fermi-Dirac statistics. It happens, however, that for high values of T the number of degenerate electrons decreases and the square curve of Fig. 135 goes over for higher temperatures to a distribution whose highest energy electrons more and more resemble in their velocity distribution the Maxwellian distribution as in the dotted portion of Fig. 135. In the limit of very high temperatures the Sommerfeld equation yields the usual Maxwellian form. What then occurs is that the faster electrons (those corresponding to the asymptotic dotted portion of Fig. 135) are the ones to escape most often, and since the average energy is μ and the total work of escape is C, the electrons escaping have an energy in excess of $C - \mu = \chi$ in the measure of the asymptotic portion of the curve of Fig. 135, which is closely Maxwellian. It is then in conformity with the new Sommerfeld theory that the thermionic emission yields a Maxwellian distribution for the relatively few electrons emitted, while inside the metal the large proportion of the electrons have the Fermi-Dirac distribution.

The values of the constants of the equation for thermionic emission of a few metals are given below.

$$i = AT^2 e^{-b/T} = AT^2 e^{-\chi/kT}$$

$$A = 1.80 \times 10^{11} \text{ E.S.U./cm}^2\text{ deg}^2 = 60.2 \text{ amp/cm}^2\text{ deg}^2$$

$$k = 1.371 \times 10^{-16} \text{ ergs/deg}$$

$$\chi_0 \text{ in volts } (\chi \text{ at } 0^\circ \text{ absolute}) = 300\frac{kb_0}{\epsilon} = 8.61 \times 10^{-5} b_0 \text{ volts}$$

RECENT THERMIONIC CONSTANTS FOR DIFFERENT METALS

Element	A	b_0	χ_0 in volts
Platinum.	1.7×10^4	72,500	6.27
Molybdenum	60.2	50,900	4.38
Tantalum	60.2	47,200	4.07
Tungsten.	60.2	52,400	4.52
Thorium	60.2	38,900	3.35
Calcium...	0.12	35,000	3.02
Nickel	26.8	32,100	2.71
Carbon	5.93	45,700	3.93

It is thus seen that by varying the temperature the saturation current can be varied at will and we can predict the approximate emission to be obtained from any filament at a given temperature if b is known.

144. THE CURRENT-POTENTIAL CHARACTERISTIC OF THERMIONIC EMISSION

It is next essential that we obtain a knowledge of how the current varies with the field applied, for in applications of the thermionic

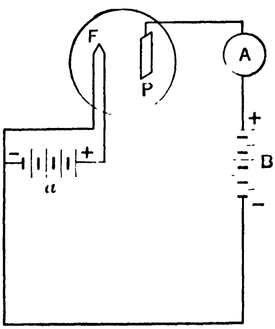


FIG. 141.—Measurement of the Current-Potential Characteristic of Thermionic Emission.

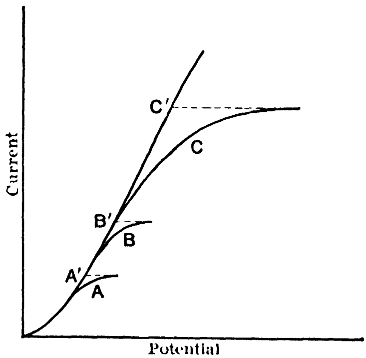


FIG. 142.—Current-Potential Characteristics of Thermionic Emission at Three Temperatures Increasing from A to C.

emission the field applied plays a most important rôle. An experimental arrangement such as shown in Fig. 141 in which the current from the filament F, or an equipotential hot cathode to a plate P, is measured by a milliammeter A or galvanometer as a function of the voltage serves for such a study. The curves obtained for different temperatures T_1, T_2 and T_3 of the cathode source are shown in Fig.

142. It is seen that the curve starts at 0, rises along what might appear to be a parabolic curve, varying as some power of the potential greater than unity, and then reaches a saturation value, or at least flattens out and approaches a saturation value gradually. This form of the curve was early recognized to be due to the fact that for heavy electron emissions with a weak field between anode and cathode the electrons emitted formed a cloud in front of the cathode which gave what is known as a *negative space charge* (i.e., a negative charge in space due to a volume distribution of negative electrons). Thus while the field draws electrons away from the cathode the accumulated negative charge of the electrons in the space between anode and cathode exerts a force driving the electrons back to the cathode and limits the current. When the applied field becomes so high that the electrons are drawn across to the anode with such speed that there is little electron accumulation between the plates (small space charge), then the maximum current possible at the temperature is acquired and we have the saturation current. In 1906, J.J. Thomson computed the limitation of space charge on the current in vacuum assuming the electrons all to emerge with the same velocity. In 1911, C. D. Child gave the solution assuming the initial velocity of emission to be zero (an assumption which is not seriously in error, for the electrons emerge with velocities equivalent to a fraction of a volt only). More rigorous equations on the basis of the real distribution of velocities are due to Langmuir and Schottky in 1913 and 1915. The derivation of Child does not give the gradual transition of the curve as saturation is approached, but this region is not of fundamental importance and for the present purposes Child's derivation and equation will suffice.

Let us assume two plates, the cathode being at 0 potential, the anode at $+E$ volts, and the distance between anode and cathode being d. The plane of the cathode passes through the origin and the distance d is measured parallel to the x axis. Now one of the important contributions of the mathematical physicist Poisson lay in setting up an equation relating potential and the volume density of the electrical charges at a given point. The equation is an expression of the continuity of flux of lines of electrical force, and the contribution made by the accumulation of charges per unit volume at the point in question. As will be remembered, the number of lines of force per square cm normal to the flux expresses the electrical field strength F. The rate of change of this field produced by a flux of lines of force due to ρ charges per cm^3 in an infinitesimal element of volume at x cm from the cathode, due to the accumulation of electrons between the

cathode and anode, is $4\pi\rho$, as there are 4π lines of force per unit charge (see page 184) and there are ρ charges per cm^3 at x. Hence $\frac{dF}{dx} = 4\pi\rho$ (see page 189). But in Chapter XV it was shown that $Fdx = dV$, where dV is the change in potential over a distance dx, hence $\frac{dF}{dx} = \frac{d^2V}{dx^2}$. We can then write Poisson's equation for our electron accumulation at a point between the plates as $\frac{d^2V}{dx^2} = 4\pi\rho$. Here it is seen that ρ must vary with x, for near the cathode the electrons will be densest, and *at* the anode there will be no electrons. At saturation of course the density will be more nearly uniform (electrons being rapidly drawn away from the cathode), but for lower potentials the electron density near the cathode will be great, falling off towards the anode. The field strength at any point x will not be $\frac{E}{d}$ but will vary in a more complicated manner as shown in Fig. 143, where curve A is the gradient $\frac{E}{d}$, while curve B shows the distortion produced by the space charge. The curve B exists for 0 velocity of electron emission. If there is a finite velocity of emission the curve at first falls below the 0 axis and then rises above, as shown by the curve C of Fig. 134. At any point x the electron has fallen through a potential V, which is the potential difference between 0 and the potential V at x. The electron has therefore obtained a kinetic energy $\frac{1}{2}mv^2 = Ve$. Finally the current density i (current per cm^2 normal to the line of flow) is given by $i = \rho v$. Here v is the electron velocity, m is the mass of the electron and e is its charge.

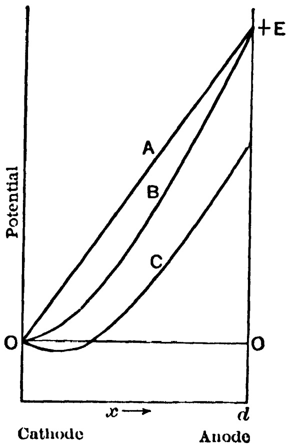


FIG. 143.—Potential Distribution between Cathode and Anode with Thermionic Emission from the Cathode; *A*. Normal Potential Drop, with No Space Charge; *B*. Drop Neglecting Initial Electron Velocities; *C*. Actual Drop.

Substituting

$$\rho = \frac{i}{v}, \quad \text{and} \quad v = \sqrt{\frac{2Ve}{m}}, \quad \text{whence} \quad \rho = \frac{i}{\sqrt{\frac{2Ve}{m}}},$$

for ρ in Poisson's equation we have

$$\frac{d^2V}{dx^2} = 4\pi i\sqrt{\frac{m}{2Ve}}$$

which on integration gives

$$\left(\frac{dV}{dx}\right)^2 - \left(\frac{dV}{dx}\right)_0^2 = 8\pi i\sqrt{\frac{2m}{e}}\,(V^{1/2} - V_0^{1/2})$$

where V_0 is the potential at the cathode (assumed 0 in this case) and $\left(\frac{dV}{dx}\right)_0$ is the potential gradient or field at the cathode. The latter is also 0 at the cathode for the case where the electrons have no initial velocity. This is seen in the curve B of Fig. 143 where the curve is asymptotic to the x axis (i.e., V has 0 rate of change with the distance x at this point). Accordingly we can write

$$\left(\frac{dV}{dx}\right)^2 = 8\pi i\sqrt{\frac{2mV}{e}}.$$

Integrating this with $V = 0$ at $x = 0$ and $V = E$ at $x = d$ (i.e., at the anode) we get

$$i = \frac{1}{9\pi}\sqrt{\frac{2e}{m}}\,\frac{E^{3/2}}{d^2}.$$

Since π, e and m are constants we can write that the space charge limited current (i.e., below saturation) is given by

$$i = 2.33 \times 10^{-6}\,\frac{E^{3/2}}{d^2}\text{ amperes per cm}^2,$$

or if the area of the cathode is A we can write

$$i_A = 2.33 \times 10^{-6}\,\frac{AE^{3/2}}{d^2}\text{ amperes.}$$

It is seen then that the current increases as the $\frac{3}{2}$ power of the potential driving the current, and inversely as the square of the distance between anode and cathode. The curve follows this law until the saturation current (A', B' and C' of Fig. 142) is reached, after which it is constant and is given by the Richardson equation depending on

the cathode temperature. The presence of the velocity distribution rounds off the otherwise sharp break of the $E^{3/2}$ curve at the saturation voltage so that saturation is much more gradually approached and the curves A, B, C of Fig. 142 take the rounded form of the experimental curves. The initial velocities also serve to prolong the rise of the asymptotic foot of the curves near the axis, giving a slower initial rise. There is still a minor correction to be used where the filament has a drop in potential due to the iR drop of the heating current. The rounded curves illustrated in Fig. 142, which contain all these factors as they represent experimental curves, are called the *characteristic curves of the two electrode tube*, the name given the *thermionic vacuum tube* with anode and hot cathode.

145. THE THERMIONIC RECTIFIER

It is at once seen that an evacuated tube with a hot cathode having a potential difference placed between cathode and anode will react in such a way as to have its current varied as the potential difference varies. If we impress an alternating current across such a tube with an electron emission from the cathode it will be seen that as long as the cathode is positive and even a little beyond this value no current will flow from anode to cathode, as the electrons are held at the cathode by a combination of space charge forces and retarding field. As soon as the cathode becomes distinctly negative the current of electrons flows to the anode. If the alternating E.M.F. rises well above the saturation value for the current a current results that rises rapidly from 0 from the beginning of the negative phase on the cathode, reaches saturation, and continues until the impressed E.M.F. drops down to zero, decreasing along the characteristic curve as the E.M.F. falls below saturation value. Hence we see that such a tube can *rectify* (i.e., give a direct current) for one-half the wave of an alternating E.M.F. This property of the 2 electrode vacuum tube is used very widely today in obtaining direct currents of high potential, from alternating potentials. The 2 electrode tube is called the *thermionic valve* or *rectifier*, and a most common form sold in this country is called a *kenotron*. To rectify both halves of the wave we can use two valves acting in inverse senses. If power is being drawn from the tube it is seen that for each valve the current is flowing half the time when there is no supply through the rectifier. Hence the potential falls during the time of the positive phase of the cathode. It is thus customary to attempt to reduce the irregularities in potential to a minimum and to smooth out the effect of the half cycle surges

by accumulating the charge on capacities of large value. The higher the frequency of the alternating potential the less serious will be the irregularities or ripples of the rectified potential, and hence it is

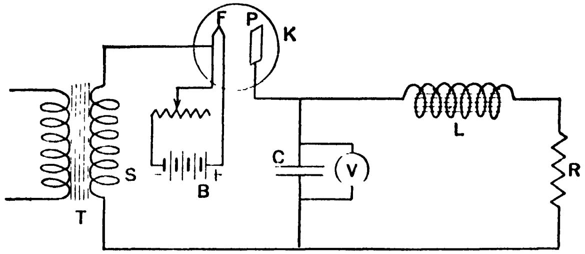


FIG. 144*a*.—Single Valve Rectifier.

deemed best to use at least a frequency of 500 cycles in the alternating potential to be rectified where smoothness of the rectified potential is required. Greater regularity in high potential direct current generation for x-ray work using valves is obtained by inserting induct-

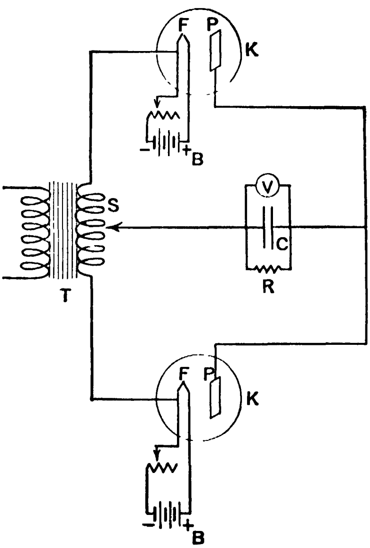


FIG. 144*b*.—Double Valve Rectifier.

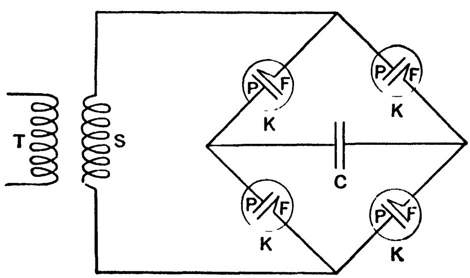


FIG. 144*c*.—Four Valve Rectifier.

ances of high value in the supply lines to " choke " out the ripples due to time variations in current. These are called choke coils. A single valve rectifier as used in rectification is shown with connections in Fig. 144*a*, while Fig. 144*b* shows a double valve rectifying both

halves of the alternating potential, and Fig. 144*c* shows a 4 valve arrangement giving even more complete rectification. The figures are self-explanatory, *F* being the filaments, *P* the plates, *K* the rectifiers, *S* the secondaries of the transformers, *T* supplying the power to be rectified, *B* the batteries heating the filaments and *V* the voltmeters measuring the rectified potential while *L* are choke coils, *C* are capacities, and *R* are resistance loads. In Fig. 145*a* the impressed and rectified potentials of the scheme of Fig. 144*a* are shown, the dotted lines being the impressed and the full lines the resultant rectified potential on *K* and *C* with a small load on *C* through the high resistance *R*. In Fig. 145*b* the currents flowing into *C* are shown compared to the impressed potential shown in the dotted curves. Figs. 146*a* and 146*b* show similar curves for the potentials on *K* and *C* and for currents in the scheme of Fig. 144*b*.

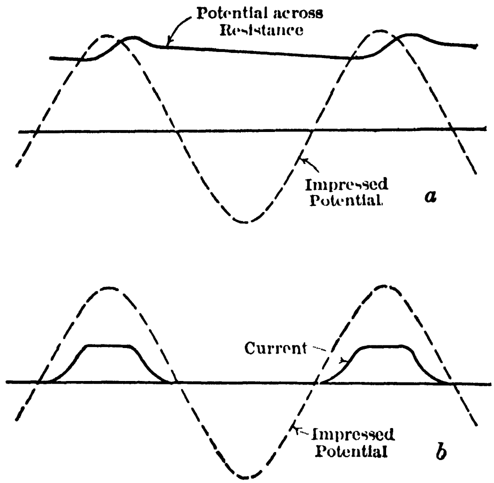


FIG. 145.—Impressed Potentials and Rectified Potentials and Currents for a Single Valve Rectifier.

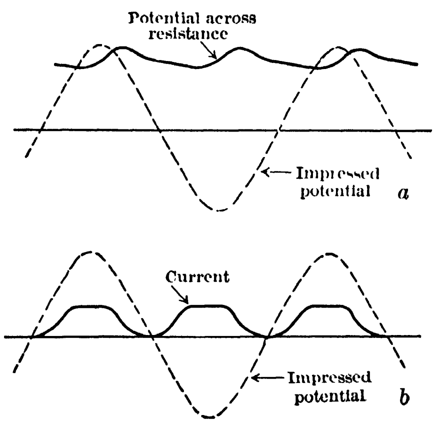


FIG. 146.—Impressed Potentials and Rectified Potentials and Currents for a Double Valve Rectifier.

These rectifiers are made for all sorts of uses from that of giving plate voltage in radio sets (B battery eliminators) to the rectifiers for thousands of volts and many amperes (power rectifiers) used in powerful radio transmitting stations, or for high potentials (10^5 volts) and currents of the order of milliamperes in x-ray work.

146. THE THREE ELECTRODE TUBE

In about 1908 a new feature was added to the thermionic rectifier valve which made it the basis of all modern developments in radio communications. It was devised as a radio detector replacing the troublesome crystal, electrolytic rectifier and the coherer by Lee DeForest. While DeForest's original device was imperfect (it contained gas) and while the inventor little understood the theory of the device it is much to his credit that he made the great step forward which once and for many years to come solved the radio communication problem. The detector was later developed as an oscillator for generating undamped electrical oscillations. DeForest called the tube the *audion detector*. The device used by DeForest had already in 1902 been used by Lenard to control photoelectric currents, and was devised in a different form by O. von Baeyer in 1908 for the control of thermionic currents. DeForest it appears, however, was the first one to exploit the device in radio detection and obtained the patent on it.

What DeForest did was to introduce a third electrode in the form of an open grid of coarse mesh between the filament and plate of an ordinary thermionic rectifier. Hence such tubes are often called 3 *electrode* tubes. The principle of the operation of the 3 electrode tube is about as follows. It was seen that in the rectifier the current through the tube was determined by the effect of the accumulated space charge due to the electrons between the anode and cathode. This was controlled to some extent by the potential between filament and plate, or anode. By placing a grid whose potential can be controlled from the outside into the space between anode and hot cathode we have the power of influencing the current through the tube by an outside agent. Were the third electrode a solid plate we would get nearly complete electrostatic screening of the anode from the cathode, and were the cathode and grid connected then the whole drop of potential would take place in the space between the third electrode and the anode so that no electrons could get through. If now the third plate be a screen of fine wires it will act to screen the cathode from the anode to a large extent. If the potential of the open grid and cathode are equal then it will be found that the anode will not be completely screened but that a few lines of force will go through the grid from anode to cathode. Hence a few electrons can wander through the grid as a result of the attenuated field, and a much reduced electron current over that in the absence of the grid from cathode to anode results. The larger the meshes of the grid the less the screening action of the grid and the greater the " transparency " of the grid

for electrons. If the grid be given a potential which is not that of the cathode, but is that which the place that it occupies would normally take on for a given current, the current to the anode would practically flow as if the grid were not there, except for a small absorption of electrons by the grid due to the area of its wires. If the grid is made more positive than the position which it occupies would normally take on, it is clear that this field between grid and cathode will *stimulate* the flow of current from the cathode by neutralizing some of the space charge or altering the space charge distribution. Hence a grid positive to the cathode by an amount greater than that of the space in the absence of the grid will increase the current to the anode, beyond that which would exist in the absence of the grid. If the grid be more negative than the place which it occupies would normally become, then it will act to repel electrons back to the cathode, thus causing an increase of the space charge at the cathode, so that the current is decreased. It is accordingly seen that owing to its strategic position between anode and cathode a grid can even by a slight negative or positive change in its potential produce a considerable change in the current through the tube, even though the meshes in the grid are very wide. The fundamental equation for the action of such a grid was worked out by Van der Bijl* in 1913, and was developed from a study of the device shown in Fig. 147. The tube *FGP* having a glowing filament *F*, a grid *G* and a plate *P*, is connected through an ammeter *A* to a battery in the plate-filament circuit of potential E_p. Between *F* and *G* there is a battery E_g putting a potential between *F* and *G*, making *G* negative to *F*. Omitting the contact potential between *F*, *G* and *P*, we can now consider the action of the grid. Due to the field between *P* and *F* at some point (say the plane where the

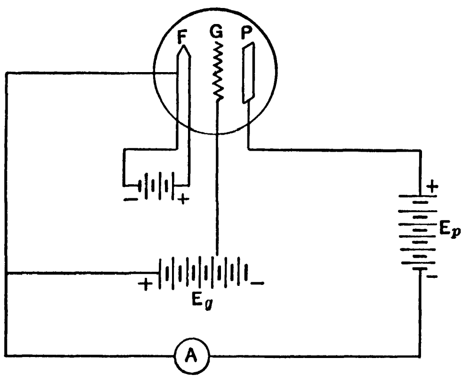


Fig. 147.—Arrangement for a Study of the Characteristic Curves of a 3 Electrode Tube.

* For a study of such problems from a physical point of view there is no text in English which excels the pioneer monograph by H. J. Van der Bijl entitled "The Thermionic Vacuum Tube," published by McGraw Hill in 1920. Despite the recent progress, the text remains the most valuable scientific introductory text on the subject for students interested more in the general physical aspects of the problem.

grid is located) there is produced a potential which can be set as the same as that of a plate of potential $\frac{E_p}{\mu}$ placed at the position of the grid, where μ *is a number* by which E_p must be divided to give the potential produced at this point by E_p. This represents the field existing in the absence of the grid due to the potential between P and F. This field as stated is positive, and acts to draw electrons from F, through the point where G is located, towards P. Now some electrons reach G, but with open meshes most of them go through the the grid towards P. Thus the current can be stimulated by replacing the field due to P by a potential of value $\frac{E_p}{\mu}$ on the grid, to nearly the extent that the current would be stimulated by E_p alone in the absence of the grid. If the maximum velocity of emission of electrons from the cathode in equivalent volts be V and a contact potential K exist between F and P, and if $K + V = \epsilon$, we can see that if a potential $-E_s$ be placed on the grid equal and opposite to that which the plane of the grid would have due to E_p, the field of E_p will be annihilated and no electron can leave the filament and reach P. Hence the negative stopping potential E_s of the grid (i.e., the negative potential applied to the grid which will stop a current due to a potential E_p on P) is given by

$$E_s = \frac{E_p}{\mu} + \epsilon.$$

Hence we see that the action of the grid on the electrons from the filament is dependent on E_g the grid potential, $\frac{E_p}{\mu}$ and ϵ. That is, the effective potential driving electrons from F to P in a 3 electrode tube is $\frac{E_p}{\mu} + E_g + \epsilon$. If $E_g = -\left(\frac{E_p}{\mu} + \epsilon\right)$ it is seen that the net result is 0 and no current will flow. If E_g is positive the current is stimulated, while if it is negative and less than $\frac{E_p}{\mu} + \epsilon$ the current is reduced. In general then we can write Van der Bijl's equation for a current i in the tube as

$$i = f\left(\frac{E_p}{\mu} + E_g + \epsilon\right),$$

where $f()$ means some function of the expression in parentheses and of

course also of the filament area, temperature, etc. The exact form of the function is uncertain but it is of an approximate form

$$i = \alpha\left(\frac{E_p}{\mu} + E_g + \epsilon\right)^2$$

and varies with the geometrical form of the tube. The factor μ was found by Van der Bijl to be a constant for a given tube and depends on the diameter of the wires and the meshes and the location of the grid between F and P. It among other things expresses the maximum amplification obtainable for a given tube and is called the amplifying factor. This equation is the fundamental equation of the thermionic amplifier.

Now the property which the third electrode imparts to the tube is that it endows the rectifier with a very sensitive sort of a trigger (the grid), which can stop (E_g negative and equal to or greater than $E_s = \frac{E_p}{\mu} + \epsilon$), or stimulate the flow of current through the tube. In this way a small potential E_g relative to the plate potential E_p, $\frac{E_g}{E_p} = \frac{1}{\mu}$ can cause a current to flow or cease flowing. Hence very feeble oscillations on G (small E_g) can be made to cause a current of much larger magnitude to flow from P to F in synchronism with themselves. By the successive building up of the currents by impressing the oscillations in the plate-filament circuit of a first tube on the grid and filament of a second tube and so on for several stages, the weakest oscillations can be amplified up to values such as to be easily converted to sound or so as to cause marked electrical effects.

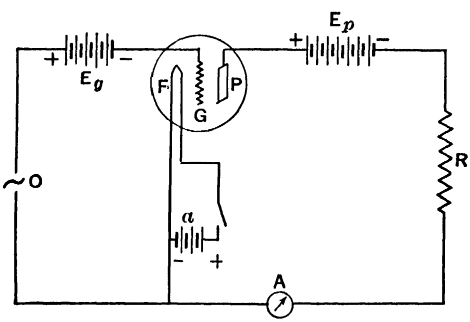


FIG. 148.—Circuit Illustrating the Amplifying Action of a 3 Electrode Tube.

A simple circuit for a study of this action is shown in Fig. 148. The oscillations entering at O are impressed between the filament F and a grid G. The filament F and the plate P are connected through a resistance R and hot wire milliammeter A to a battery E_p. a is the filament battery rendering the filament incandescent, and E_g is a bias battery for varying the potential of G negatively to F in order to let

the oscillations impressed on the grid work over a particular part of the range of action of the grid on the current i between P and F. With a negative E_g on G equal to or greater than E_s there will be no current through R. Under other conditions a current flows from P to F (electrons from F to P), and with oscillations of sufficient amplitude on the grid an intermittent current flows through R, giving an intermittent iR drop in R in synchronism with the oscillations in O. This iR drop can be impressed between filament and grid of a second tube, or else R can be replaced by a coil coupled inductively* to a coil in a second circuit. If the oscillator be omitted one can study the

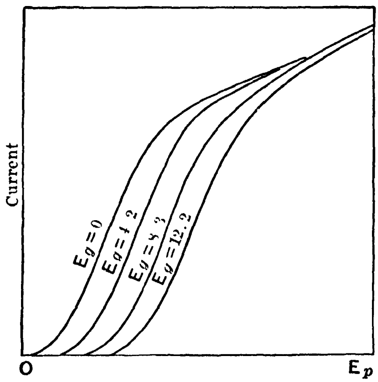


FIG. 149.—Characteristic Curves for 3 Electrode Tube, Current Plotted against *Plate Voltage* for Different Grid Potentials.

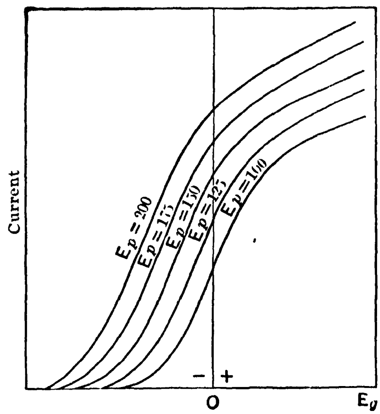


FIG. 150.—Characteristic Curves for 3 Electrode Tube, Current Plotted against *Grid Potential* for Various Plate Voltages.

current i to the anode or plate as a function of E_p and of E_g. This gives one a set of curves for different values of the two independent variables. Such a set of curves are reproduced in Figs. 149 and 150, which result from the measurements of Van der Bijl.† In Fig. 149 the current is plotted against E_p for different values of E_g, and in Fig. 150, the current is plotted against E_g for different values of E_p. In Fig. 150 it is seen that the value of E_g required to reduce the current to 0 is the more negative the greater E_p. The curves are the *characteristic curves* of the 3 *electrode tube.*

* The amplified oscillations in the branch $FARE_rP$ can be impressed between the grid and filament of a second tube by means of either the resistance method or else by inductive coupling, the conditions of operation determining the method used.

† These curves were taken from Van der Bijl's book "The Thermionic Vacuum Tube," and are reproduced through the courtesy of the McGraw-Hill Publishing Co.

If it is desired to rectify the incoming oscillations it is seen that when these oscillations pass through the 0 phase, the current should just be 0 (i.e., the bias voltage E_g should have the value shown in Fig. 150 for the foot of the curve corresponding to the E_p used). Then on the negative half of the wave no current will pass. During the whole positive phase the current will flow, starting at zero and rising along the characteristic curve up to the maximum current drawn by the peak positive potential impressed on G by the oscillations, and then falling off to 0 along the curve, the current curves showing much the same form as in the two electrode rectifier, Fig. 145*b*. If on the other hand it is merely desired to *amplify* the oscillations positive and negative, *without distortion*, the *straight portion of the characteristic alone must be used.* In this case it is assumed that the *voltage amplitude* of the incoming waves is *small* compared to the *range along the x axis of the curves of Fig.* 150 over which the slope of the curves is constant. Take for example an E_p of 200 in Fig. 150. The curve is nearly linear between an E_g of 23 volts and 8 volts. The oscillations must then have a total amplitude of less than 15 volts. E_g must accordingly be set so that the 0 value of the impressed E.M.F. lies at about 15 volts. Then positive and negative oscillations will cause changes in current which are proportional to the potential variations. These variations in current can be made to give potential variations of the order of μ times the incoming amplitudes, and we have an amplifier.

In general, radio signals consist of a train of sine wave form high frequency oscillations whose intensity varies in a fashion dependent on the vibrations impressed on the sending oscillator by the voice through a microphone or by a telegraph key. These oscillations are transmitted as a wave train through space and reach an antenna or collecting aerial wire of large absorbing power tuned to resonance with the frequency of the incoming oscillations. These are communicated by means of a radio transformer (having no iron) to an amplifying tube or series of such tubes. The amplified oscillations are thus intensified to a point where they can be rectified, and are converted by a rectifying 3 electrode tube to unidirectional surges of varying intensity and duration. If the transmitted signal is a sound signal the integrated unidirectional surges have frequencies so low (10,000 cycles or less) that they permit of amplification by iron core transformers, hence they are once more placed on an audio frequency 3 electrode tube amplifier capable of a high energy output. This in turn acts on the coils of a sound generator or loud-speaking microphone. The myriads of devices used to amplify and reproduce such signals make it impossible in a text of this scope to give any

particular circuits or class of circuits, so that the principles only are given at this point. Furthermore the types of tubes and the methods of achieving various goals are changing so rapidly that it hardly pays to describe circuits that are out of date before a book is off the press.

147. THE THERMIONIC OSCILLATOR

Not long after the development of the 3 electrode tube as an amplifier and rectifier its use for an even more important purpose was discovered. As was seen in Chapter XXV a capacity connected to an inductance which naturally has some resistance will, if the electrical state of the circuit be disturbed, lead to an electrical oscillation. The period T of the latter is given approximately by the equation $T = 2\pi\sqrt{LC}$, where L is the self-induction and C is the capacity. Owing to the resistance of the circuit such oscillations are highly damped and are useless for any but telegraphic communication (i.e., they cannot be modulated by the voice). The damping also reduces the range of distances over which signals may be sent. Various attempts were made to produce sustained (i.e., nearly undamped) electrical oscillations of high frequency. Among these the successful attempts led to the design of generators of 100,000 cycles or more of relatively small energy output, and of high frequency oscillating arcs such as the Poulson and Chaffee arcs. The latter were relatively successful and gave large power outputs with sustained oscillations, which were however incapable of being used for anything but telegraphy as the intensity of the oscillations could not be controlled or modulated by sound waves, etc. The 3 electrode tube however solved the problem of furnishing sustained oscillations which are capable of modulation and led at once to the development of radiotelephony and television, as well as to increasing the range of radio communication in general.

The exact manner in which the oscillating power of these tubes was discovered is somewhat obscure. There is no one name associated with this discovery in the literature and two facts probably served to make it difficult to trace its origin. One of these is the fact that this phase was developed for commercial uses and hence the secrecy surrounding patentable inventions was a contributing cause to the slow dissemination of information. Secondly much of the development occurred during the World War, 1913 to 1919, under cooperative and needless to say secret investigations which were divulged only after 1919 in a completed form. The fact that the output energy of a 3 electrode tube amplifier is greater than the input energy, and that

hence a part of the output energy might be fed back to sustain the oscillations of the input portion of the system, was doubtless the idea that must have struck many workers and led to the development of the 3 electrode tube oscillator. The period of the oscillation of such a device will depend, as do the periods in any oscillating circuits, primarily on the values of self-induction and capacity in the circuit, and hence these act as the " timing clocks " in such oscillators. The operation of one or two types of oscillators will be indicated below in order to show the nature of this use of the 3 electrode tube, though no attempt will be made to discuss the quantitative theory.

An excellent example of a simple oscillating circuit is the Hartley circuit pictured in Fig. 151. A capacity C is connected to two inductances L_1 and L_2 in series. The circuit L_1L_2C, in which C may be variable, having self-induction capacity and resistance (of the coils), will oscillate electrically if disturbed as given by the equations of Section 127 (i.e., this is the " time clock " of the circuit). A battery B is placed in series with the plate and F through L_2. As the current from B starts to flow through L_2 when the filament F is heated, the potential drop across L_2 starts a current charging C through L_1. The system CL_1L_2 then oscillates with a period T approximately given by $T = 2\pi\sqrt{C(L_1 + L_2)}$ if the resistance of L_1 and L_2 is small. The oscillation sets up an alternating E.M.F. between F and G which is aided by the bias battery b placing G on the sensitive part of its characteristic curve. The value of L_1 relative to L_2 is so chosen that the E.M.F. between F and G is in the proper phase relation with the oscillation in the circuit CL_1L_2 to give a sustained oscillation. Thus as the potential between F and G causes the current through the tube from P to F to vary, this same current *reinforces* the oscillations in the condenser system CL_1L_2. In this manner continuous undamped oscillations occur in CL_1L_2 of any desired frequency depending on L_1, L_2 and C, which can be transferred to any desired system by coupling coils or other devices. The test for the presence of oscillations and adjustment of the circuit is facilitated by the use of a thermogalvanometer T in the circuit CL_1L_2. The above circuit is sometimes modi-

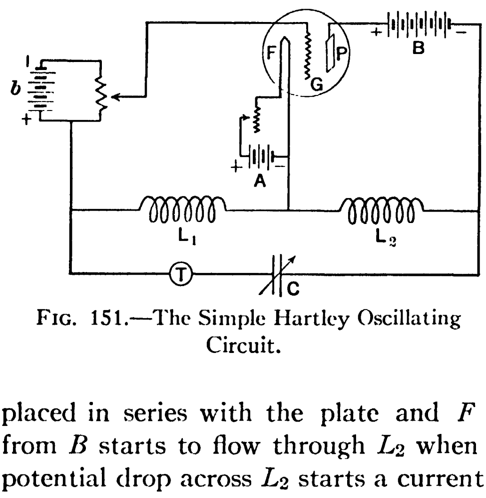


FIG. 151.—The Simple Hartley Oscillating Circuit.

fied by replacing the bias battery b by a capacity-grid leak system for charging the grid to a proper potential. This arrangement is shown in Fig. 152. The circuit is the same except that b is replaced by the condenser C_s and the resistance R_s. The resistance R_s is of rather high value (order of a megohm) and C_s and R_s are so chosen that the accumulations of electrons on G due to the electron current from F to P keeps G at the proper potential with respect to F. C_s may vary from 0.0005 microfarad to 0.002 microfarad or more. Colpitts used a circuit in which the *capacity is divided*, to play the rôle of interrupter and oscillator instead of the divided inductances as used above in the Hartley circuit. A type of Colpitts' circuit is shown in Fig. 153. In this circuit the oscillating circuit, or " timing clock " is the circuit C_1C_2L. The capacity is $\frac{C_1C_2}{C_1 + C_2}$, and the period T is approximately

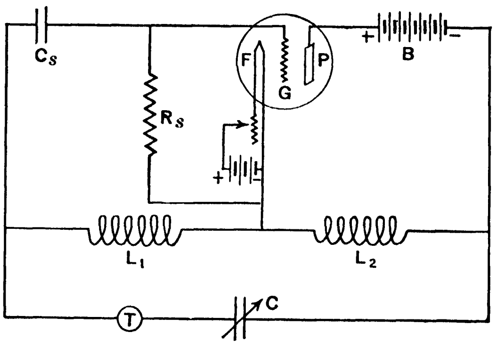


FIG. 152.—The Simple Hartley Oscillating Circuit Using a Grid-Leak Instead of a Bias Battery.

$T = 2\pi\sqrt{L\left(\frac{C_1C_2}{C_1 + C_2}\right)}$. The driving battery B drives the electron current through F to P except as the oscillations on G modify it. The current runs through an iron-cored choke coil Ch_2 which acts to damp the radio frequency oscillations through the battery. Ch_1 an iron-cored high impedance choke coil replaces the grid resistance used in the Hartley circuit, while C_3 fulfills the same function as in the latter circuit. C is a very high capacity which establishes a high frequency connection between B and P, and the condenser system C_1C_2L. The oscillations are im-

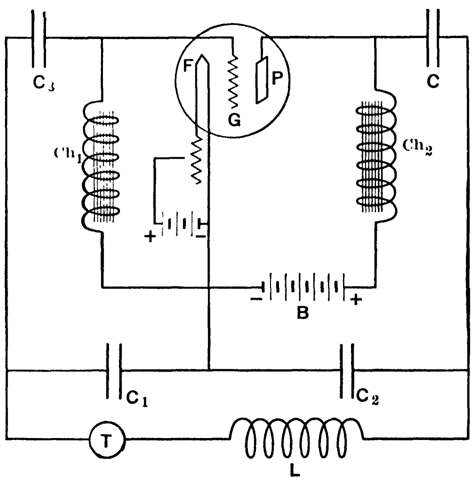


FIG. 153.—The Simple Colpitts' Circuit.

pressed on the grid by being picked off across C_1 which is in series with C_2.

A third type of oscillation generator is shown in Fig. 154. In this case the oscillating circuit is CL_2, T being the thermogalvanometer. The oscillations in CL_2 are placed between G and F through the bias battery b. The driving battery B operates to drive electrons from F to P through an inductance L_1 closely coupled to L_2. In this case the period is influenced by CL_2 and the mutual inductance M between L_1 and L_2. As soon as the plate current begins to flow through L_1 an oscillation is induced in L_2C which maintains such a phase relation to the current from B that sustained oscillations are generated in L_1 which may be picked up by a circuit tuned to L_1 as desired. The energy of oscillation in CL_2 comes from the inductive coupling with L_1 and from the battery B.

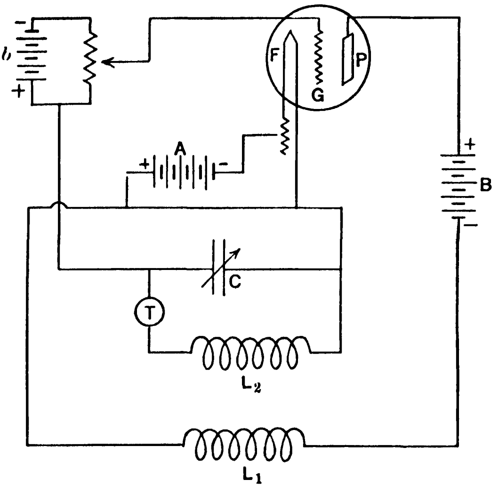


FIG. 154.—Tuned Grid Circuit Oscillator of the Hartley Type.

The variations of these circuits and the manipulation of the oscillations so obtained by amplification, rectification, etc., are indefinitely great. There is thus no place in this text for a further discussion of the application of the thermionic emission to the modern developments of applied science, sufficient theoretical consideration having been given to lay a foundation for further study.

148. PIEZO ELECTRIC AND MAGNETOSTRICTIVE OSCILLATORS

Perhaps a word or two might be said about two more sets of electrical phenomena which are now widely used in the work on oscillating systems. One of these makes use of the so-called piezo electric (pressure-electrical) effect in quartz first studied by Pierre Curie. If a quartz crystal be cut so that two of its faces are perpendicular to the optic axis, and if the faces be rendered conducting by placing metal plates thereon, or by sputtering with metal, it will be observed that whenever the crystal is subjected to mechanical stress (i.e.,

either compressed or elongated along its optic axis) an electrical displacement takes place in the crystal, one electrode becoming positive and the other negative. The charges on the plates on compression are reversed on tension, and hence by *compressing and expanding the crystal an alternating potential difference* is set up on the plates which can be amplified to any desired magnitude by thermionic amplifiers. Now such a quartz crystal will have a *mechanical* frequency of vibration dependent on the thickness of the crystal and its elastic constants. A similar behavior is manifested by crystals of Rochelle salts. Hence it is seen that by making the crystal vibrate mechanically, electrical oscillation of great constancy can be obtained. We thus have a new sort of " timing clock " (the mechanical vibrations of crystals), which can be used in place of the usual electrical oscillators. The value of these crystals, however, lies in a further property. If properly timed electrical charges are placed on the crystal faces by means of the plates there will be mechanical vibrations introduced on the quartz crystal. Hence these crystals when in tune with an electrical oscillation can be made to vibrate quite violently. The frequencies of these vibrations are of course well below the customary radio frequencies, but lie well above the frequencies of audible sounds (i.e., between ten thousand and some hundreds of thousands of cycles per second). They are therefore called *supersonic vibrations.* The energies of the supersonic waves obtained are prodigious, as the mechanical efficiency of the crystals is high and the electrical energy input is almost entirely converted to sound. Furthermore these supersonic waves of short wave length are not diffracted by ordinary objects like ships, but cast sharp shadows and travel in nearly straight lines. If an oscillator and a parabolic reflector send out a beam of such waves these will be reflected back by any solid bodies (hulls of ships in water, or the ocean bottom). The echo or reflected wave can then be picked up by a tuned crystal amplifier converted to direct current surges of the modulated frequency and heard on a telephone. From the velocity of sound and the time elapsed between sending and receiving a signal the distance of the reflecting body can be measured. These supersonic transmitters and receivers served as the most successful of the devices for detecting submarines in the World War, and have now been installed in many ships for automatically continually determining and recording the depth of water in the form of instruments known as *fathometers.* The development of supersonic devices is largely due to P. Langevin, who developed the first submarine detectors. The mechanical effects produced by the concentrated energy in supersonic waves are such as to cause intense local heating, to

break up emulsions and to cause chemical reactions. There is no doubt a great future in the application of these crystals to many problems. Among other problems it appears that the constant frequency of these crystals may revolutionize the problems of timing and time keeping, and investigations are well under way for utilizing this phenomenon in such problems as those of astronomical clocks.

The use of a similar mechanical effect of *electromagnetic* nature is also being developed today. A cylindrical bar of iron of a certain length and diameter has its own mechanical periods of oscillation. It can vibrate in tune to its natural frequency or multiples thereof. G. W. Pierce at Harvard has made use of the magnetostrictive effects of magnetizing currents in a solenoid about such a bar, to set the bar into oscillations in its natural period. These mechanical vibrations of a bar produced by the magnetostrictive effects of currents of proper frequency can be transmitted to a sounding board of adequate design. Hence by properly tuning a bar to resonance with a radio signal of appropriate frequency the modulations impressed on this radio frequency by a microphone adequately amplified can be impressed on the amplitude of vibration of the bar. In this way the sounding board can be made to reproduce the audible notes impressed on the microphone with very great precision. What the future of such magnetostrictive devices in practical application will be remains for the future to show.

PROBLEMS BASED ON CHAPTERS III AND IV

1. In the diagram of Fig. P1, A is a magnet whose length is 10 cm. Find the magnitude and direction of the resultant field at a point P, 10 cm distant from the N pole of A on a line at right angles to A, if the pole strength m is 40 units. This problem is to be solved numerically.

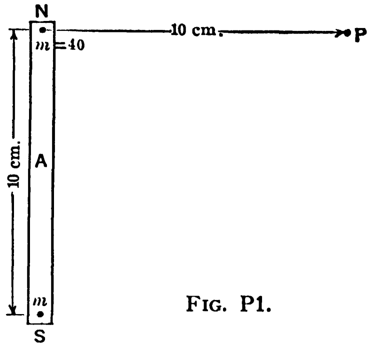


FIG. P1.

2. Given a magnet whose length is 16 cm and whose pole strength is 200 units. Calculate the magnitude (in gauss) and direction of the resultant field at a point distant 12 cm from the south pole and 20 cm from the north pole of the magnet.

3. The magnet A of the Fig. P2 is 5 cms long and has a pole strength of 100 units. It is placed at right angles to the earth's magnetic field H of strength 0.25 gauss as

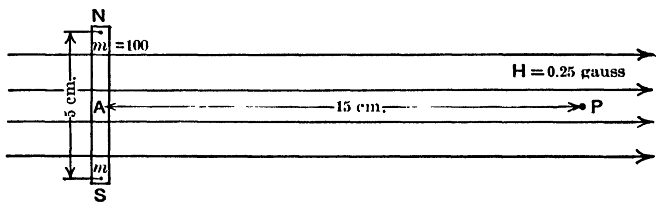


FIG. P2.

indicated by the arrows. Calculate the resultant field at a point P distant 15 cm from the center of A on a line at right angles to the axis of A, in magnitude and direction. In this problem use the approximate formula for the strength of the field at a point on the perpendicular to the center of a magnet.

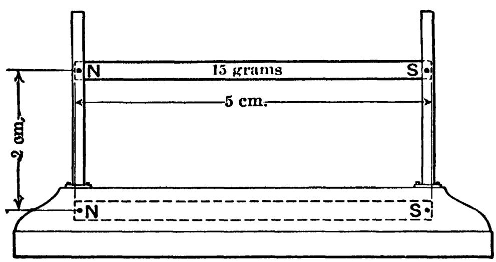
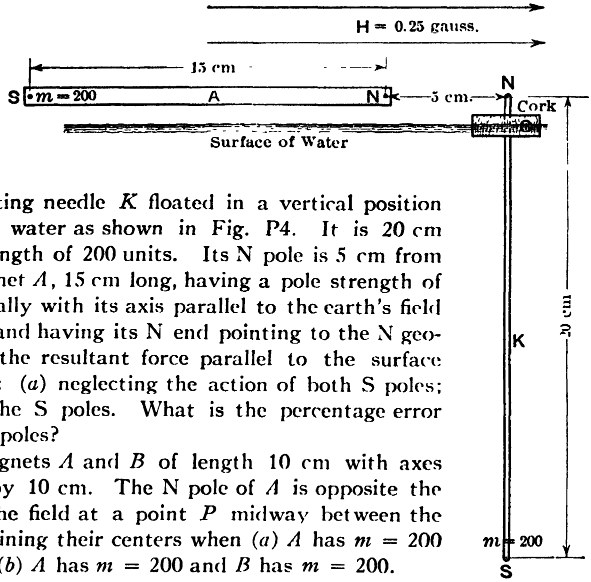


FIG. P3.

4. In the lecture demonstration experiment a small bar magnet 5 cm long weighing 15 grams was shown to be supported with its axis 2 cm from an identical bar magnet concealed in the wooden box below it as shown in Fig. P3. The N pole of the suspended magnet is just over the N pole of the magnet in the box. These magnets are made of a new alloy developed by the Western Electric Company and represent examples of as powerful a magnet as can be obtained for small

permanent magnets of this type. Calculate the pole strength of these magnets from the data given above, taking into account attractive forces as well as repulsive forces.



5. Given a steel knitting needle K floated in a vertical position by means of a cork C on water as shown in Fig. P4. It is 20 cm long, and has a pole strength of 200 units. Its N pole is 5 cm from the N end of a bar magnet A, 15 cm long, having a pole strength of 200 units, lying horizontally with its axis parallel to the earth's field H of strength 0.25 gauss and having its N end pointing to the N geographic pole. Calculate the resultant force parallel to the surface of the water for 2 cases: (a) neglecting the action of both S poles; (b) taking into account the S poles. What is the percentage error due to a neglect of the S poles?

6. Given two bar magnets A and B of length 10 cm with axes parallel and separated by 10 cm. The N pole of A is opposite the S pole of B. Calculate the field at a point P midway between the two magnets on a line joining their centers when (a) A has $m = 200$ and B has $m = 100$, and (b) A has $m = 200$ and B has $m = 200$.

FIG. P4.

7. Calculate the field due to the magnet A in magnitude and direction at a point P located as indicated in the diagram of Fig. P5 if the magnet has a pole strength of 100 units and is 20 cm long.

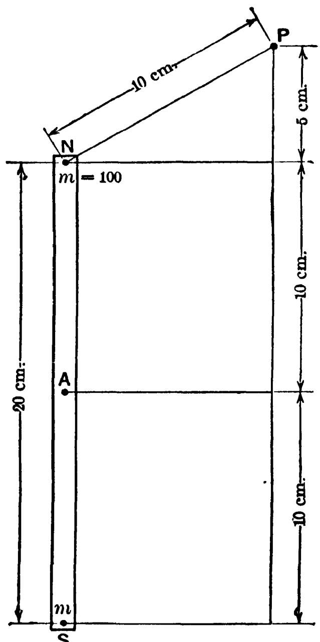


FIG. P5.

8. Calculate the resultant field in magnitude and direction at a point P on the axis of the magnet A, Fig. P6, of 50 units pole strength and 10 cm length, 50 cm from its center at its north end, and 50 cm from the center of a magnet B which lies on the axis of A but is at right angles to this axis and has a length of 10 cm and a pole

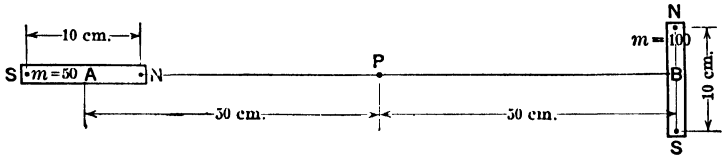


FIG. P6.

strength of 100 units, its north end being upward. This answer need not be given more accurately than 5%.

9. A bar of copper is suspended in a stirrup by a fine steel wire and its position accurately set perpendicular to the horizontal component of earth's field H. The bar is then removed and replaced by a powerful bar magnet of magnetic moment $M = 4 \times 10^4$ units. This magnet sets itself so that it makes an angle of 30° with its initial position and 60° with the earth's field. The radius of the wire is 0.030 cm and its coefficient of rigidity $\eta = 5 \times 10^{11}$. The length of the wire is 38.4 cm. Calculate the horizontal component H of the earth's magnetic field to an accuracy not exceeding 3 significant figures. Remember that the angles used in the formulae are in radians.

10. A bar of copper is suspended by a fine steel wire by means of a stirrup in a position perpendicular with the earth's field H whose horizontal component is at this place 0.15 gauss. When the copper bar is removed and a magnet of moment $M = 6.78 \times 10^4$ units is placed in the stirrup the magnet takes on a position in the field such that the wire and stirrup are twisted through φ degrees with the perpendicular to the field (the original rest position), and with $90 - \varphi$ degrees or $\theta°$ with the field. If the radius of the wire is 0.03 cm, its coefficient of rigidity is 4×10^{11} and its length is 30.7 cm, calculate the angle φ which the magnet makes with the perpendicular to the earth's field. Do not carry computations further than 3 significant figures. (Note when the equation is properly set up an expression of the form, $\dfrac{\cos \varphi}{\varphi} = A$, results in which A has a numerical value. To solve such an equation plot the values of $\dfrac{\cos \varphi}{\varphi}$ as ordinates for various angles φ from 0 to π radians, against φ as abscissae. At the point A on the plot read off the value of φ, which is the desired result, and convert to degrees.)

11. Referring to Problem 9, calculate through what angle the torsion head of the suspension would have to be twisted to hold the bar magnet at an angle of 80° with the field, using the torsional constant given in Problem 9 and the value of the earth's field there calculated of 0.254 gauss. (To solve this problem remember that there is at 80° with the field a torque on the magnet; this torque must be equal to the torque on the suspension produced by twisting the torsion head.)

12. Problems 12, 13, and 14 use the same data given under I and II.

I. A bar magnet of unknown moment M is placed with its axis horizontal and perpendicular to the horizontal component of earth's field H which is also unknown. A small compass needle at 31.1 cm from the center of the bar magnet along its axis is observed to be deflected 45°, with the earth's field.

II. When this magnet M was mounted on a stirrup suspended by a fiber of negligible rigidity its period of oscillation in the earth's field was found to be π seconds. When the same magnet suspended as above was loaded with a ring whose moment of inertia was 450 g cm² the period was observed to be 2π seconds. The approximate length of the magnet was 10 cm.

Calculate the ratio $\frac{M}{H}$ for the magnet.

13. Calculate the product MH for the magnet from the data of Problem 12.

14. Determine the values of M and H separately from the data given in Problem 12, and calculate the approximate pole strength m of the magnet. If the angle of dip is 70° calculate the H_T total intensity of the earth's field.

15. Place a short bar magnet whose magnetic moment is 500 in any position you choose with its axis in a horizontal plane in which there is a uniform field of 0.20 gauss. Having placed your magnet in this field and indicated its position locate quantitatively all positions of neutral points on the paper; that is, points where the field intensity is 0, using approximate formulae.

16. Problems 16, 17 and 18 use the same data given under I and II.

I. A bar magnet of unknown moment M is placed with its axis horizontal and perpendicular to the horizontal component of the earth's field H which is also unknown. A small compass needle at 28 cm from the center of the bar magnet along its axis is observed to be deflected 60° with the earth's field.

II. When this magnet M was mounted on a stirrup suspended by a fiber of negligible rigidity its period of oscillation in the earth's field was found to be π seconds. When the same magnet suspended as above was loaded with a ring whose moment of inertia was 360 g. cm², the period was observed to be 2π seconds. The approximate length of the magnet was 5 cm. Calculate the product MH for the magnet and field.

17. Calculate the ratio M/H for the magnet from the data in Problem 16, part I.

18. Determine the values of M and H separately from the data given in Problem 16, and calculate the approximate pole strength m of the magnet. If the angle of dip is 75° calculate the H_T total intensity of the earth's field.

PROBLEMS BASED ON CHAPTERS V AND VI

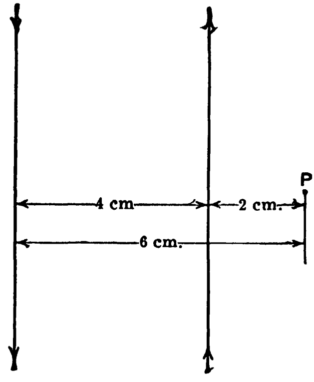


FIG. P7.

1. Given two long straight parallel wires, such as might feed a motor as indicated in Fig. P7, separated by 4 cm. The current flows up in one and down in the other. If the current carried by them is the same and equal to 10 amperes calculate the resultant field H in gauss at a point P, 2 cm from one and 6 cm from the other, and indicate the direction of the resultant field.

2. A student was measuring the field at a point P distant 40 cm from the center of a circular coil A along its axis. The coil had 100 turns and a radius of 30 cm carrying a current of 1 ampere in the sense indicated by the arrows of Fig. P8. Accidentally a wire B ran vertically upward in the wall of the room 20 cm from P. If it

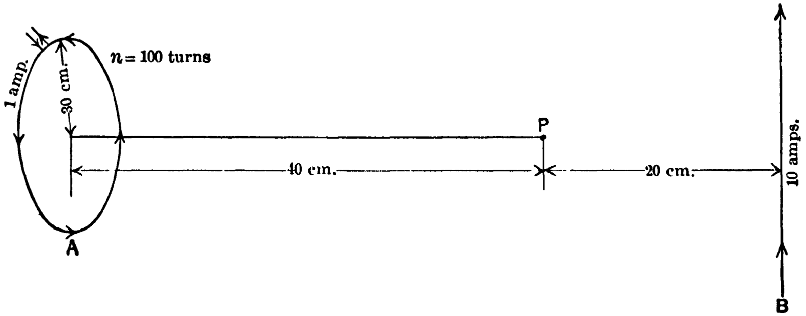


FIG. P8.

carried a current of 10 amperes flowing upwards calculate in magnitude and direction the fields he measured on two separate occasions: (*a*) when no current was flowing in B; and (*b*) when the current was flowing in B.

3. In a laboratory experiment a student was investigating the field due to a circular coil of 10 turns, radius 15 cm, at a point P 50 cm from its center along its axis. The current in the coil was 2 amperes and in the direction of the arrows in Fig. P9. He measured the field by observing the deflection of a compass needle in the field

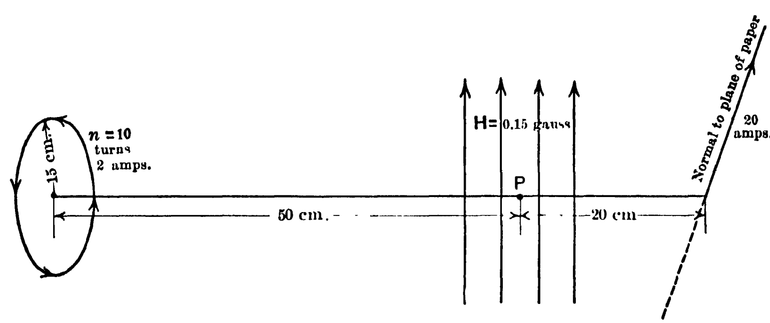


FIG. P9.

of the coil and the horizontal component of the earth's field which was one of 0.15 gauss and at right angles to the axis of the coil. Unknown to the student there was in the wall 20 cm distant a vertical power supply line wire carrying a current of 20 amperes upwards. Calculate the angle θ of the compass needle with the earth's field and the resultant field: (*a*) when the power line was running (indicate direction on diagram); (*b*) when the power was off.

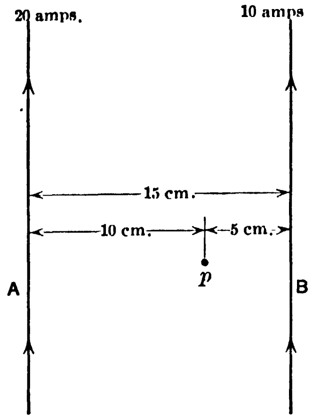


FIG. P10.

4. Given two long parallel wires A and B shown in Fig. P10 15 cm apart with currents of 20 and 10 amperes respectively flowing upwards in them. Calculate the resulting field in magnitude and direction at a point p, 10 cm distant from A and 5 cm distant from B.

5. Given the two coils A and B shown in Fig. P11, having a radius of 20 cm and 100 turns each. A current of 5 amperes flows in coil A and 2 amperes flows in coil B

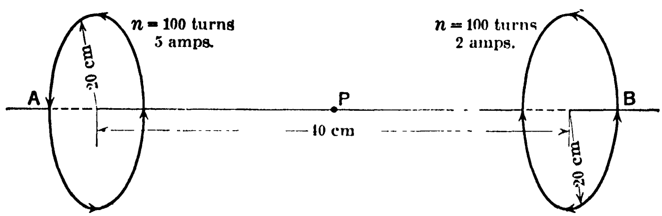


FIG. P11.

in the direction of the arrows. The coils have the same axis and are separated by 40 cm. What is the direction and magnitude of the field at a point P on their common axis 20 cm from each coil?

6. In the earth's field of horizontal intensity 0.25 gauss a suspended magnet had a period of 2 seconds. The same magnet placed in the center of a plane circular coil of radius 15.7 cm having 75 turns so placed that its field adds to the earth's field is observed to have a period of 1 second. Calculate the current in the coil in amperes.

7. A current passed through a tangent galvanometer, and through a silver voltameter connected in series. In the voltameter, metallic silver was deposited at the cathode. The galvanometer had 30 turns a radius of 15.70 cm, the horizontal component of the earth's field H was 0.12 gauss and a deflection of 45° was observed. This same current in 2 hours deposited 0.8050 gram of silver. (*a*) How many grams of silver does one coulomb deposit? (*b*) How many coulombs deposit 107.9 grams of silver?

8. In testing magnetos for aeroplane engines during the war all the characteristics of the sparks were required to be known. To test the heat per spark and thus the average ignition current during the spark, a standard spark gap was placed inside a heat insulated copper block calorimeter and the rate of heating of the block was determined as the magneto operated. The heating curve with the magneto attached

giving 360 sparks per minute was compared to calibration curves taken with a heating coil placed in the block, the current through which, and potential across which, could be measured. With a given magneto the curve coincided with that due to a current of 0.5 ampere at 120 volts. The average voltage across the spark gap was known from oscillograph measurements to be about 200 volts (it started at 6000 volts and dropped to 100 volts during the discharge so as to give an average of 200 volts). The time of the discharge was known by oscillograph to be 0.001 second. (*a*) Calculate the watts delivered to the copper calorimeter by the heating coil and therefore by the 360 sparks per minute. (*b*) Calculate the average rate of temperature rise of the copper if its equivalent mass was 1000 grams (specific heat .093), before it began to lose much heat by radiation (i.e., on the straight part of its heating curve). (*c*) Calculate the power delivered in watts per spark from the magneto. (*d*) Then calculate the average current from this magneto in amperes during the discharge of 0.001 second.

9. An unknown P.D. was applied to a resistance wire in a calorimeter whose water equivalent was 528 gram calories per degree C. The current was run for 3 minutes and 20 seconds and the temperature rise amounted to 10° C with a current of 1 ampere. What was the potential difference in volts across the calorimeter, and what was the power consumption in watts?

10. A lightning discharge struck a copper lightning rod 1 cm in diameter and 1000 cm long. The density of copper is 8.5, its specific resistance is 2×10^{-6}, its melting point is 1100°, its latent heat of fusion is 42 calories per gram, and its specific heat is 0.12. The rod was fused and the heat of the discharge went to heating it to its melting point and fused it. The discharge lasted 10^{-5} seconds. With these data calculate:

(*a*) The weight of copper melted.

(*b*) The heat liberated in the process.

(*c*) The resistance of the lightning rod.

(*d*) The average current in amperes.

(*e*) The potential across the wire if the discharge was constant for the 10^{-5} seconds.

(*f*) The power liberated in watts and the quantity of electricity in coulombs.

(*g*) The magnetic field in gauss 10 meters from the rod.

(*h*) If the flash which struck the rod was 300 m long in the air and the P.D. was 10^9 volts the following data are required:

(1) The power in watts liberated in the air.

(2) The apparent resistance of the air path of the spark as a whole and the resistance per meter.

11. An electric flatiron is wound for a 110-volt circuit and takes 5.5 amperes of current under operating conditions. What is its power consumption? How many calories of heat are developed per hour? What does it cost to operate the iron at the rate of 5 cents per kilowatt hour?

12. A coil of resistance wire 528 cm long is connected to a 110-volt power main and is observed to raise the temperature of a liter of water in a beaker 5 degrees C in 1 minute and 40 seconds. (*a*) What is the resistance of the wire? (*b*) If the wire had been 5 times as long, what would the rise in temperature have been in the same time?

13. A short compass needle free to vibrate in a horizontal plane is suspended 10 cm magnetically east of a long vertical wire. The horizontal component of the earth's field at the compass needle is 0.25 gauss. When there is no current in the wire the needle vibrates 30 times a minute. A current i is sent through the wire and it is

observed that the needle reverses direction and now makes 20 vibrations per minute. Determine i in amperes and state whether i is up or down in the wire.

PROBLEMS BASED ON CHAPTERS VII AND VIII

1. Determine the resistance between the points A and B in the circuit shown in Fig. P12. The numerals denote the resistance of each conductor in ohms.

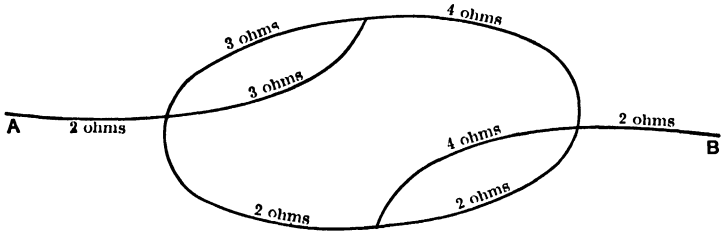


FIG. P12.

2. A wire of 10 ohms resistance is drawn out until its length is doubled. Assuming that the specific resistance remains constant calculate its resistance.

3. Copper has a specific resistance of 1.8 times 10^{-6}. A telegraph line of 100 km length is to be run from one town to the next. The length of the wire will thus be 200 km of copper wire of *diameter* 0.356 cm (about No. 10 wire). It requires 0.319 ampere to work the relay at the receiving end. (*a*) How many Daniell cells of E.M.F. = 1.15 volts each will be required to operate the line? (*b*) If the temperature ranges from $-20°$ C in winter to $+40°$ C in summer, and if the temperature coefficient of resistance of copper is 0.004 ohm, per ohm per degree C, how many cells will be required to care for this change, and at what time of the year must more cells be used?

4. Telegraph lines are run for given stretches of less than a hundred miles each. This is because the resistance of the wire reduces the current. (*a*) If it takes 0.159 ampere to work the relay at the farther end of a stretch, and if it is convenient because of losses to use only 115 volts on a line, how many relay stations will have to be used from here to Chicago 3200 km (twice the length of wire is used for the current and its return) if the wire is copper of diameter 0.356 cm (about No. 10) whose specific resistance is 1.8×10^{-6}. (*b*) If the temperature ranges from $-30°$ C to $+40°$ C from winter to summer, and if the temperature coefficient of resistance for copper is 0.004 ohm per ohm per degree C, what will be the potentials needed on the line in the coldest and hottest weather if 115 volts was needed at 0° C?

5. A battery has an E.M.F. of 20 volts and an internal resistance of 2 ohms. An ammeter in series with a resistance is connected to the battery terminals and the ammeter reads 4 amperes. A wire of 3 ohms resistance is now connected to the battery terminals in parallel with the ammeter branch. What will the ammeter now read and what will be the difference in potential between the battery terminals?

6. A voltmeter of very high resistance is connected to the terminals of a battery whose internal resistance is 3 ohms and reads 18 volts. If an ammeter joined in series with a resistance giving a total resistance of 6 ohms be connected to the terminals of the battery in parallel with the voltmeter what will each instrument read?

7. Given a voltmeter of ranges 150 volts and 15 volts. For the 150-volt scale

the resistance is 300 ohms, and for the 15-volt scale it is 20 ohms. When the potential across 100 dry cells of E.M.F. of 1.5 volts each connected in series was read the 150-volt scale read 144 volts. Ten cells in series were then put on the 15-volt scale. (*a*) What was the internal resistance of 1 cell? (*b*) What did the 15-volt scale read?

8. A small generator gives a potential reading on a voltmeter of 100 volts when the voltmeter of resistance 300 ohms is placed across a resistance of 100 ohms, which is in series with the generator. When the 100-ohm resistance is replaced by a 200-ohm resistance the voltmeter reads 105 volts. Calculate the internal resistance R_i and the E.M.F., E of the generator. (Note two simultaneous equations in R_i and E can be obtained from the data above.)

9. The circuit depicted in Fig. P13 is typical of the circuits used in wiring houses. In order to choose the wiring to be used it is essential to know the currents in each branch. (*a*) Calculate the currents i_1, i_2, i_3, i_4, and i_5 as well as the potential across

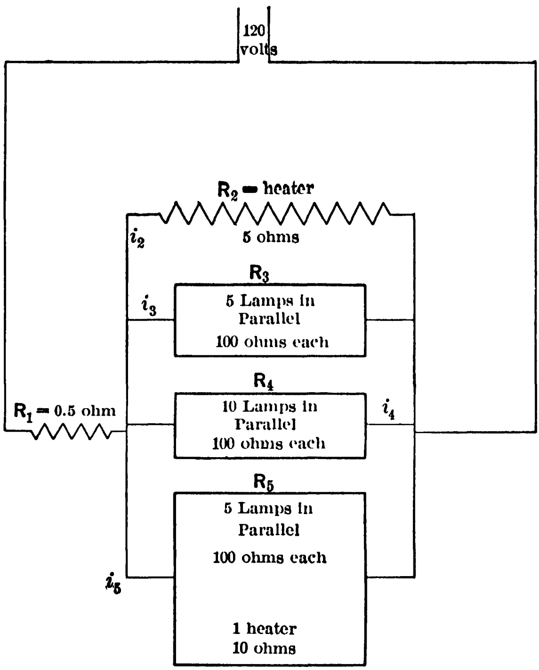


FIG. P13.

R_1, R_2, R_3, R_4, R_5 when the lights are all burning. (*b*) Calculate the potentials across the resistances R_1 and R_2, R_3, R_4, R_5 which are in parallel when the currents in both heaters are 0, that is, with the circuits R_2, and 10 ohms in R_5 cut out. The change found in potential across the lamps shows why it is that lights dim when electric heaters are put on.

10. A cell of internal resistance R_i and E.M.F., E gives 2 volts potential across a resistance of 20 ohms, and 1.5 volts potential difference across a resistance of 10 ohms. Calculate R_i and E.

Problems on Kirchhoff's Laws of Divided Circuits

11. Given a Wheatstone's bridge network which is in "balance" ($i_g = 0$). Prove that the battery and galvanometer may be interchanged without destroying the balance.

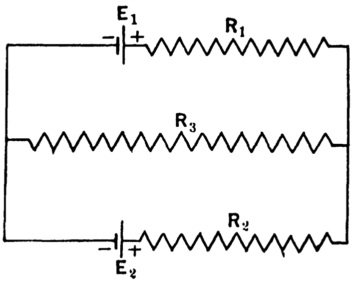


FIG. P14.

12. Given the circuit shown in Fig. P14 with the following values of potentials and resistances. Calculate the currents i_1, i_2 and i_3 in amperes, in fractional form, using Kirchhoff's laws.

$E_1 = 3$ volts
$E_2 = 2$ volts
$R_1 = 2$ ohms
$R_2 = 1$ ohm
$R_3 = 2$ ohms

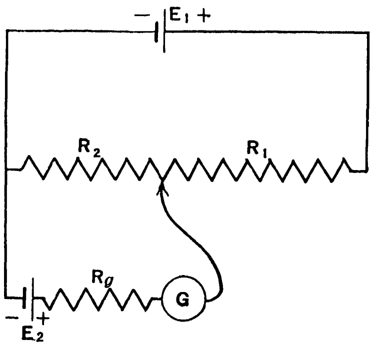


FIG. P15.

13. Given the circuit shown in Fig. P15 which is a potentiometer circuit. Calculate the currents in each of the three branches of the circuit. Given:

$E_1 = 2.0$ volts
$E_2 = 1.0$ volt
$R_1 = 60$ ohms
$R_2 = 50$ ohms
$R_g = 50$ ohms

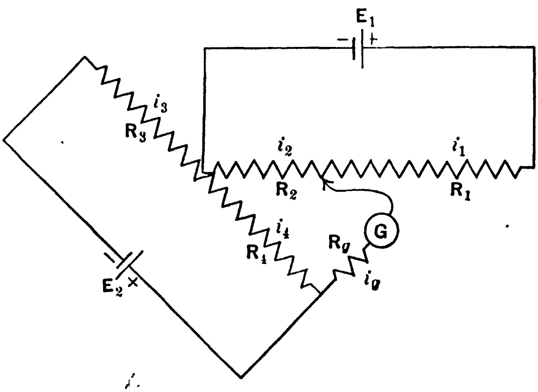


FIG. P16.

14. Given the circuit shown in Fig. P16, solve for the currents i_1, i_2, i_3, i_4 and i_g. Given:

$E_1 = 5$ volts
$E_2 = 1$ volt
$R_1 = 10$ ohms
$R_2 = 5$ ohms
$R_3 = 5$ ohms
$R_4 = 10$ ohms
$R_g = 10$ ohms

15. Given the circuit shown in Fig. P17, which represents a Wheatstone bridge out of balance. If the galvanometer can stand a current of 10^{-4} amperes will this lack of balance ruin it? Determine the answer by solving for i_g only. Given:

$E = 1$ volt
$R_e = 1$ ohm
$R_g = 40$ ohms
$R_x = 20$ ohms
$R_R = 30$ ohms
$R_P = 40$ ohms
$R_Q = 50$ ohms

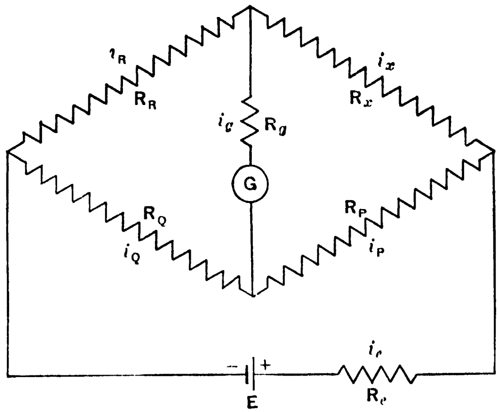


FIG. P17.

16. Given the circuit shown in Fig. P18. It is a Wheatstone Bridge out of balance. Solve for the value of i_e, i_x, i_R, i_g, i_P, and i_Q in amperes (*to be given in fractional form*). If a current of 10^{-3} amperes would wreck a galvanometer that it was intended to use, could such an unbalance as this be safely used with the galvanometer? Given:

$R_R = 100$ ohms
$R_x = 80$ ohms
$R_g = 100$ ohms
$R_P = 100$ ohms
$R_Q = 100$ ohms
$R_e = 1$ ohm
$E = 1$ volt

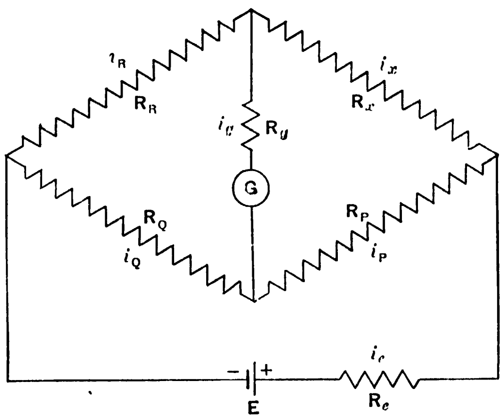


FIG. P18.

17. Given the circuit shown in Fig. P19. Solve for the currents i_1, i_2, i_3, i_4, and i_g in amperes in fractional form. Check solution by substituting in one of the other equations. Given:

$R_1 = 30$ ohms
$R_2 = 15$ ohms
$R_3 = 5$ ohms
$R_4 = 10$ ohms
$R_g = 100$ ohms
$E_1 = 10$ volts
$E_2 = 10$ volts

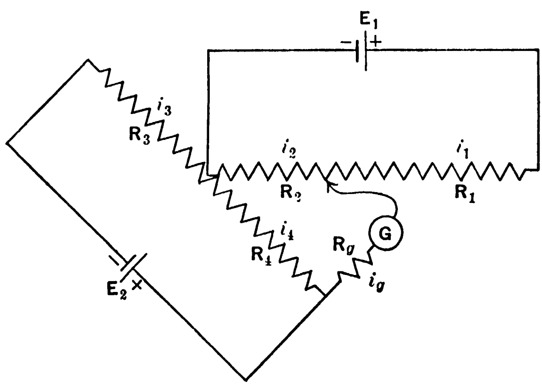


FIG. P19.

PROBLEMS BASED ON CHAPTERS IX, XI, XII AND XIII

1. Given a D'Arsonval galvanometer built according to the following specifications: It has a coil of 100 turns, length 8 cm, width 0.75 cm. The field H is one of 500 gauss. The phosphor bronze suspension has a modulus of rigidity $\eta = 6 \times 10^{11}$ dynes/cm^2, whose radius is 4×10^{-3} cm and whose length is 2π cm. The moment of inertia of the coil is 154 dynes cm^2. Calculate (*a*) the galvanometer constant K; (*b*) the figure of merit k, (*c*) its sensitivity in megohms; (*d*) its period of oscillation. All the dimensions are realizable. Do you consider this a practical galvanometer? Can you see how a much greater sensitivity may become impractical with a galvanometer like this?

2. A milliammeter has a resistance of 1 ohm and a scale reading from 0 to 100 milliamperes. (*a*) What is the resistance of a shunt that would convert it to an instrument reading from 0 to 10 amperes? (*b*) What series resistance would one use to convert it to a voltmeter reading from 0 to 10 volts?

3. A laboratory voltmeter reads from 0 to 150 volts. Its resistance is 300 ohms. It is required to use it for measuring potentials from 500 to 600 volts. Show what resistances are needed and how they are to be connected to attain this result enabling one to use the full range of the scale and a convenient multiplying factor.

4. An ammeter has a range of from 0 to 1 ampere. It is needed to measure currents from 15 to 25 amperes. Its resistance as measured was 1.0 ohm. (*a*) How would you connect it in a circuit to read these currents? (*b*) What resistances if any would you use? (*c*) How would you deduce from them the true current through the circuit?

5. Given a D'Arsonval galvanometer made according to the following specifications: It has a coil of 200 turns, length 6 cm, width 0.5 cm. The field H is 400 gauss. It has a phosphor bronze suspension of modulus of rigidity $\eta = 6 \times 10^{11}$ dynes/cm^2 whose radius is 5×10^{-3} cm and whose length is π cm. The moment of inertia of the coil is 77 dynes cm^2. Calculate: (*a*) The galvanometer constant; (*b*) the figure of merit; (*c*) its sensitivity in megohms; (*d*) its period of oscillation. Can you see why a much greater sensitivity might become impractical in a galvanometer like this?

6. J. J. Thomson measured the deflection of a stream of electrons in a magnetic and electric field and from the measurements found the ratio of the charge e (the quantity of electricity) on an electron to the mass m to be given by $e/m = 1.81 \times 10^7$ absolute electromagnetic units of quantity per gram mass of electrons. Now the mass of an electron is 1/1860 that of the H atom or 8.8×10^{-28} grams. Calculate from these data the quantity of electricity in absolute electromagnetic units carried by the electron. What is the value in coulombs? If each chlorine ion carried as its negative charge 1 electron, and there were 6.06×10^{23} atoms of chlorine in a gram atom, how much electricity would a gram atom of chlorine ions carry? How does this compare with the quantity actually carried by a gram atom of chlorine ions and what does it lead one to conclude about the nature of the charge carried by the chlorine ion?

7. Given that a gram atom of oxygen has in it 6.06×10^{23} atoms. It takes 193,000 coulombs to liberate this gram atom of divalent oxygen atoms. (*a*) What charge does the divalent oxygen atom carry in coulombs, in E.M.U.? (*b*) The mass of the hydrogen atom is 1.662×10^{-24} grams, and that of an electron is 1/1860 that of the H atom. Calculate the ratio of charge to mass, e/m, in electro-magnetic units per gram carried by the electrons assuming that their charge is one-half that on the

oxygen ion. Compare this with the value of $e/m = 1.81 \times 10^7$ absolute E.M.U. per gram observed by J. J. Thomson for cathode rays and draw your conclusions.

8. A water pipe takes part of the return current from a trolley line. The current is spread over an area of the pipe amounting to 100 cm by 20 cm. The current is 0.08 ampere, the ground being negative to the pipe. If the pipe is 0.5 cm thick how long will it take to corrode the pipe away, provided the corrosion is uniform over this area and assuming in the moist ground away from air that the Fe goes into solution as Fe^{++}, or ferrous iron.

9. Ten liters of pure H_2 gas are required at 27° C and 760 mm pressure for an experiment. The H_2 is to be generated electrically and must be generated in a period of less than 5 hours. What current will be required if 96,500 coulombs deposit 1 gram of H_2 gas?

10. A small lead storage cell is charged with a current of 1 ampere. The cell is such that 25 grams of water if taken from it will lower the level of the acid in the cell 1 cm. The cells were left to charge overnight at the above mentioned current. At 5 A.M. the cells were fully charged and the current was used in decomposing the water. How far had the level fallen by 10 A.M. when the current was shut off?

11. When Na reacts with Cl to form NaCl, 97,900 calories of heat are given out for every gram atom of Na combining. If it takes 96,500 coulombs to liberate 23 grams of Na what will be the E.M.F. of a cell with Na and Cl electrodes, neglecting entropy?

12. In a Daniell cell $ZnSO_4$ is formed and $CuSO_4$ is decomposed depositing Cu on the anode. The heat of formation of $ZnSO_4$ is 248,000 calories per gram-molecule. That of $CuSO_4$ is 197,500 calories per gram-molecule. From these data calculate the heat set free per gram-molecule of the reaction of the cell, and then per gram Zn consumed. If Zn has an atomic weight 65.4 and a valency 2, calculate the E.M.F. of the Daniell cell in volts. Given $J = 4.18 \times 10^7$ ergs.

13. The Daniell cell composed of a cell in which Zn replaces Cu in $CuSO_4$ has a potential of 1.05 volts. The heat of formation of $CuSO_4$ is 197,500 calories per gram-molecule. Given the atomic weight of Zn = 65.4, its valency as 2, and J as 4.18 $\times 10^7$ ergs calculate the difference in heat of formation of $ZnSO_4$, and $CuSO_4$ first per gram and then per gram-molecule, and hence the heat of formation of $ZnSO_4$.

14. A piece of apparatus must be plated with pure gold to resist the action of acids. The deposit must be 0.1 mm thick and the chamber is a cylinder of 5.62 cm radius and 20 cm high. The inside of the cylinder and the base must be plated. The gold is trivalent and the current density must not exceed 0.005 ampere per cm^2 of gold surface. How many hours will it take to plate the chamber?

15. In the thermoscope shown in Fig. P20 the current through the system is 20 amperes. The resistance of the Cu — Bi element in the small bulb of 75 cm^3 capacity is 5.0×10^{-4} ohms in each cell. The Peltier coefficient $P = 10^{-2}$ for each element and the current flowed for 147 seconds. During this period there was a quantity of heat developed in each of the two junctions. This heat must be calculated for each junction. Since the heat capacity of the 43 grams of metal rod (Bi + Cu) heated, together with the small heat capacity of the gas in each bulb, was 2.59 calories for a degree rise in temperature, calculate the temperature rise observed in each bulb. Then calculate the difference of level in the arms of the manometer of each bulb; given the density of the liquid in each column equal to 1, and the outside pressure as 76 cm of mercury, assuming the air in the bulb at the temperature of the metal, and

the room temperature as 20° C. The density of mercury is taken as 13.6. (*Note:* Remember that in one junction there is a Peltier heating while in the other there is a Peltier cooling, as well as a Joule heating in both junctions, and the two

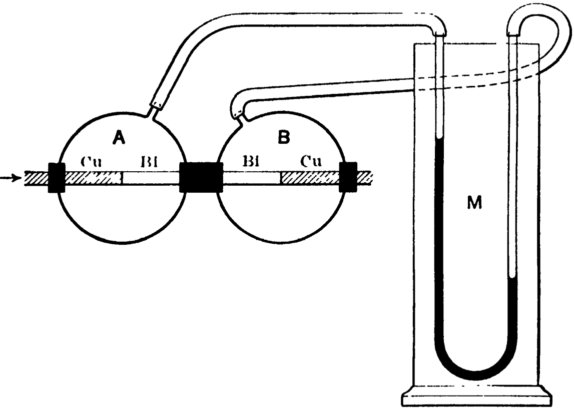


FIG. P20.

effects add in one cell and subtract in the other. Do not carry out computations to more than two significant figures as the data do not warrant greater accuracy.)

16. The thermoelectric power of iron is $(17.2 - 0.048t) \times 10^{-6}$ volts; that of copper is $(1.34 + 0.0094t) \times 10^{-6}$ volts. Calculate the thermoelectromotive force for an Fe – Cu couple when one junction is kept at 0° C by ice while the other takes on the temperature 50°, 100°, 150°, 200°, 250°, 300°, 350°, 400°, 500° and 600°, and plot them. The calculation is to be carried out graphically and should not be accurate to more than three significant figures. What is the temperature of the neutral point?

17. The thermoelectric power of iron is $(17.2 - 0.048t) \times 10^{-6}$ volts; that of Ni is $(-21.8 - 0.05t) \times 10^{-6}$ volts. If one junction is kept at 0° C while the other takes on the temperatures 50°, 100°, 150°, 200°, 250°, 300°, plot the values of the thermoelectromotive force observed. Then give two reasons by comparing this couple with those in Problems 16 and 18 why the Ni–Fe couple is admirably suited for work between 0° and 300° C.

18. Given the thermoelectric power for iron as $(17 - 0.05t) \times 10^{-6}$ volts and for aluminum as $(-0.8 + 0.004t) \times 10^{-6}$ volts, calculate the thermoelectromotive force of an Al–Fe couple when one junction is at 0° C and the other junction is at 50°, 100°, 150°, 200°, 250°, 300°, 350°, 400°, and 500° C. What is the temperature of the neutral point? The calculation is to be done graphically and accurate to not more than three significant figures, accompanied by plot.

19. Given a set of 500 small lead storage cells such as used in radio work. They have an E.M.F. of 2.2 volts each and an internal resistance of 0.5 ohm. They are to be used in a circuit of 10 ohms resistance. What is the most efficient arrangement of these cells and what current will they give?

20. In the thermoscope shown in Fig. P20 the current through the system is 25 amperes. The resistance of the Cu–Bi element in the small bulb is 4.0×10^{-4} ohms in each cell. The Peltier coefficient $P = 10^{-2}$ for each element, and the current flowed for 180 seconds. During this period there was a quantity of heat developed

in each of the two junctions. This heat must be calculated for each junction. Since the heat capacity of the 43 grams of metal rod (Bi + Cu) heated, together with the small heat capacity of the gas in each bulb, was 2.59 calories for a degree rise in temperature, calculate the temperature rise observed in each bulb. Then calculate the difference of level in the arms of the manometer of each bulb. Given the density of the liquid in each column equal to 1, and the outside pressure as 76 cm of mercury, assuming the air in the bulb at the temperature of the metal, and the room temperature as 20° C. The density of mercury is taken as 13.6.

PROBLEMS BASED ON CHAPTERS XIV, XV AND XVI

1. A, B and C are the vertices of an equilateral triangle, the radius of the circumscribing circle being 10 cm. A charge of +10 E.S.U. is placed at A, a charge of +10 E.S.U. at B, and a charge of −5 E.S.U. at C. Find the force in magnitude and direction on a charge of +2 E.S.U. placed at the center of the circumscribing circle.

2. Two identical conducting pith balls of mass 0.51 gram each are suspended by threads 40 cm. long. On being charged and brought in contact they take positions with their centers 20 cm apart. What is the charge on each?

3. If the pith balls in Problem 2 had been immersed in toluene of dielectric constant 2.5 what would the distance between them have been? Assume $g = 980$ dynes and carry out computations to 2 figures only. *Suggestions*: the unknown is $r = \frac{1}{2}$ the distance between balls. Two equations between force and r can be set up from which a cubic equation in r results. Solve this equation by successive approximations or graphically.

4. Two light, conducting spheres of mass 1.02 grams each are charged electrically by touching them to an electrical machine and then together. When in equilibrium they are separated by 30 cm while their suspending threads are 45 cm long. What is the charge on each, assuming $g = 980$?

5. Compute the force between an electron in a hydrogen atom and the nucleus if they are point charges of -4.77×10^{-10} E.S.U. and $+4.77 \times 10^{-10}$ E.S.U. and separated by 1×10^{-8} cm. In the K ring of electrons of uranium having a charge of $+92 \times 4.77 \times 10^{-10}$ E.S.U. the K electron having a charge of -4.77×10^{-10} E.S.U. is as close as 5×10^{-11} cm from the nucleus. Calculate the force in this case. Then realizing that the force acts on a body which can be considered as having a radius 3.5×10^{-13} cm calculate the pressure which the electron would exert in these two fields if it could press on some body.

6. When a swift α particle, a doubly positively charged helium atom of mass 6.64×10^{-24} grams, is shot against a massive nucleus of an atom of gold, which can be considered as fixed, it approaches to within 2×10^{-12} cm before its velocity is annihilated and it reverses its direction and separates. The charge on the gold nucleus is $+79 \times 4.77 \times 10^{-10}$ E.S.U. and that on the α particle is $+2 \times 4.77 \times 10^{-10}$ E.S.U. Since at the point where the α particle starts to retrace its path the potential energy equals the kinetic energy which it initially had, one can with the above data calculate the potential energy of the α particle at the point of closest approach and hence know the kinetic energy of the particle. Determine the kinetic energy of the α particle and calculate its velocity in cm/sec.

7. Calculate the energy in ergs which an electron of charge -4.77×10^{-10} E.S.U. acquires in falling through a potential difference of 1 volt, 1000 volts, 200,000 volts, and 10^6 volts (the latter potentials have recently been achieved in an x-ray tube at the California Institute of Technology).

8. Given the mass of the electron as 8.99×10^{-28} grams and assuming that mass does not appreciably change with these velocities calculate the velocities of the electrons in cm/sec from the data of problem 7. Note that the velocities obtained in the case 10^6 volt electrons exceed 3×10^{10} cm/sec. This is impossible according to relativity and we see that the increase in m should have been included.

9. A small conducting sphere of radius 1 cm was charged to +6000 *volts* and immersed in nitrobenzol, $D = 40$. A small sphere with an unknown charge having a mass of 2 grams was found to be in equilibrium with the first sphere when rotating freely about the first sphere with a speed of 1 cm/sec, in an orbit of 5 cm radius. Given the law of centripetal force as $f = mv^2/r$, calculate the unknown charge in magnitude and tell what its sign is.

10. Assume that in a solution 2 ions of opposite signs are separated by 10^{-6} cm (this would nearly be true in a gram-molecular solution of NaCl). Calculate the force on the ions first in the absence of water (i.e., in empty space or air) and then in the presence of the water, dielectric constant 81, if the charge on the ions is 4.8×10^{-10} E.S.U. of quantity. Then calculate the forces between ions at 10^{-3} cm, at 10^{-5} cm, and at 3×10^{-8} cm when the two ions are in contact. (This is about the distance between the sodium and chlorine ions in an NaCl crystal.)

11. An electric charge of +2000 units is placed on a conducting sphere of radius 5 cm. The same charge is later placed on a body having the same capacity, but of such a shape that one of the ends has a radius of curvature of 1 cm. The potential remains constant. Find the following quantities:

(*a*) The potential of the sphere in E.S.U., in volts, and in absolute E.M.U.

(*b*) Calculate the charge density on the sphere.

(*c*) Calculate the charge density on the curved end of the second body.

(*d*) How many lines of force per cm² emerge at the surface of the sphere?

(*e*) How many lines of force per cm² emerge at the surface of the sharply pointed body?

(*f*) What is the field strength close to the surface of the sphere, in absolute units/cm and in volts/cm?

(*g*) What is the field strength close to the surface of the pointed body?

(*h*) If air can stand a field strength of 30,000 volts per cm before it breaks down, will the air break down in either case?

(*i*) What would be the force on a unit positive charge 1 cm distant from the surface of the 5 cm sphere?

12. An absolute electrometer has plates 10 cm in diameter separated by 2 mm. It is desired to measure potentials of 3000 volts as accurately as possible. If weights of 0.001 gram are all that it is convenient to use in the instrument, and taking $g = 980$ dynes, how accurately can the potentials of this amount be measured?

13. An absolute electrometer has plates of 20 cm diameter separated by 0.5 mm. It is placed across a *DC* source of potential and 100 grams are needed to balance the applied field. If one gram of weight = 980 dynes, calculate the P.D. in E.S.U. A standard voltmeter gave the potential across these mains as 1300 volts. Calculate the value of the ratio of the E.S.U. of potential to the E.M.U. of potential.

14. Given the radius of the earth as 6500 km, and that of the moon as 3400 km. Assume that, by the action of the ultraviolet light of the sun on the surface of the moon which is unprotected by an atmosphere, the number of electrons that can leave

the moon is sufficiently great to give it a potential of +300 volts. The charge on the electron is 4.8×10^{-10} E.S.U. of quantity. With these data calculate:

(*a*) The capacity of the earth and moon in cm, in absolute E.M.U. in farads and microfarads, assuming the earth and moon each isolated in space.

(*b*) Assuming the moon's surface conducting calculate how much electricity would escape to raise it to 300 volts potential in E.S.U., coulombs and in absolute E.M.U.

(*c*) If the moon were to lose the above charge how much energy would be liberated?

15. The earth's radius is 6370 km, and that of the sun is 1,395,000 km. The temperature of the sun is 6000° C, which corresponds to the energy which an electron would get by falling through a potential of 0.34 volt. Thus electrons on the sun due to their heat motions can leave until the sun acquires a positive potential of this value. Given the charge of an electron as 4.8×10^{-10} E.S.U. Calculate:

(*a*) The capacity of the earth and sun in cms, farads, and absolute E.M.U. if they are considered completely isolated in space.

(*b*) Calculate the quantity of electricity the sun would lose in this way in absolute E.S.U., in coulombs, and the number of electrons which would leave the sun due to their heat motions.

16. An electrically charged cloud 1 square km in extent approached to within 100 m from the earth's surface when it discharged to a copper lightning conductor of resistance 3×10^{-2} ohms melting 400 grams of copper in the 10^{-4} seconds of the current flow and discharge. The latent heat of fusion of copper is 43 calories per gram and the copper was raised from 0° C to its melting point 1100° C with a specific heat 0.08. Calculate:

(*a*) The capacity of the cloud earth system in cms and in farads from the dimensions given, assuming the cloud and earth to form a parallel plate condenser.

(*b*) The current which flowed for 10^{-4} seconds and the quantity of electricity transferred.

(*c*) The potential between the cloud and the earth and from this the sparking potential of a lightning discharge in volts/cm.

17. A condenser such as is used in telephone circuits having a capacity of 1 microfarad was charged to 100 volts. It was then connected to a condenser of 2 microfarads capacity which was at 0 potential. Calculate the energy first in the charged condenser and then in the system of the two condensers. How much loss of energy was there, and how much were the connections raised in temperature, if they weighed 2 grams and had a specific heat 0.08 calorie.

18. Given a 100,000 volt power line. The station has a 20,000-volt static voltmeter of capacity 1000 cm. It has also a standard plate condenser that can take 20,000 volts whose capacity is 4000 cm. Calculate the dimensions of a parallel plate condenser, whose plates to stand the potential must be at least 4 cm apart, such that the voltmeter can be used, and what the multiplying factor is. Accompany the answer by a diagram showing how the voltmeter will be fixed across the line.

19. A parallel plate condenser is required capable of withstanding 10,000 volts

and having a capacity of 0.1 microfarad. Glass plates 3 mm thick immersed in oil giving a dielectric constant 6 will suffice. The glass plates available are such that the conducting surfaces of thin copper can be 40 × 30 cm. Calculate how many sheets will be needed. (Note in building up a condenser like this if there are n sheets of copper $n - 1$ condenser units will be obtained. Show that this statement is correct.)

20. A water tank having a base area of 1 m^2 is filled to a height of 2 m of water. It is allowed to fall through narrow pipes to an empty tank of the same base area until the pressure equalizes. One hundred "navy jars," having a capacity of 0.002 microfarad each in parallel, are charged to 3000 volts. They are then connected in parallel to an exactly similar uncharged condenser system. Given $g = 980$. Calculate:

(*a*) The energy of the water in the full tank and in the two tanks. The loss of energy in equalization, and state what became of the energy.

(*b*) The energy of the charged condenser systems before and after they are combined. Where did the energy difference go?

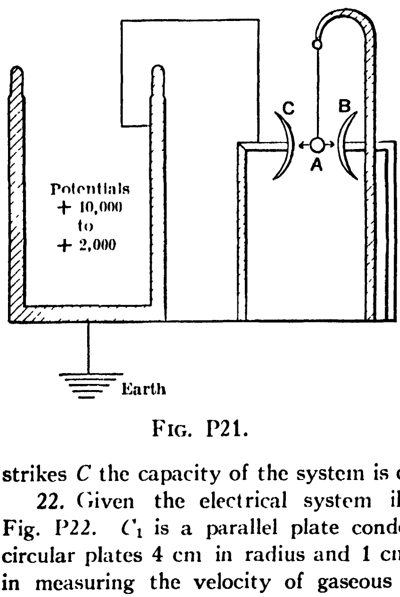


Fig. P21.

21. Given the first electric bell dating from 1750 illustrated in Fig. P21. In this the small insulated sphere A of radius 1 cm after being touched to the inner terminal bell-shaped plate C of the condenser of capacity 0.001 microfarad, which has a positive potential of 10,000 volts relative to the bell B, and the outer layer of the condenser which is earthed, is repelled until it strikes the bell B and is discharged. It is then attracted to C until it makes contact and repeats the process. The bell will keep ringing for a long time. If the forces are insufficient for the operation of the bell when the potential falls to 2000 volts, calculate the number of times the bell rings. (*Note:* When the sphere A strikes C the capacity of the system is changed and the potential is altered.)

22. Given the electrical system illustrated in Fig. P22. C_1 is a parallel plate condenser having circular plates 4 cm in radius and 1 cm apart used in measuring the velocity of gaseous ions. C_2 is a cylindrical air condenser of capacity 300 cm., and E is a source of potential of 100 volts. Calculate:

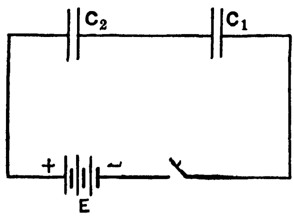


Fig. P22.

(*a*) the total capacity of the system.

(*b*) the potential across C_2 and C_1, for it is essential that the potential across C_1 be accurately known.

23. In a study of discharge through gases ions are created by radioactive deposits. The currents are measured by the potential to which a given capacity is raised. In some such experiments the capacity used is 300 cm. The potential to which it is

raised in 10 seconds is 0.1 volt. Calculate the ion current in E.S.U. and in amperes. Could a galvanometer be used for this work? If the charge on an ion is 4.77×10^{-10} E.S.U., how many ions per second constitute this current, if they are all of one sign. If each α particle produced 3.5×10^4 ions, how many α particles are emitted from the source per second, creating ions which reach the measuring field.

24. The ionization produced by cosmic rays at the surface of the earth per cm^3 per second is about 2. In a vessel of 1 liter these ions are all swept out by a field. The electrode and gold leaf system has a capacity of 2 cm. How many volts does the potential on the gold leaf fall in 5 hours? Calculate the current measured by this device in absolute E.S.U. and in amperes. How does it compare with sensitive galvanometers? Given the charge on an ion as 4.8×10^{-10} E.S.U.

25. In a study of currents in gases ions are created by radioactive radiations and by cosmic rays. The electroscope used in measurement has an electrical capacity of 2 cm. Over a period of 1 hour the ions of both signs produced by the radiations in a volume of 500 cm^3 cause a fall of 3 volts in the potential of the gold-leaf electroscope. Calculate the ion current in E.S.U. and in amperes. Could a galvanometer be used in this work? As ions of both signs take part in the transport of current and each ion carries 4.77×10^{-10} E.S.U. of charge, calculate how many ion pairs were found per second. How many ion pairs are generated per cm^3 per second?

26. In a study of the velocity of gaseous ions, ions are generated by polonium and by means of an auxiliary field and gauze the ions of one sign are selected. In 10 seconds time the current of positive ions charges a cylindrical condenser of length 30 cms, radii of inner and outer cylinders 2.5 and 2.6 cm, with air as a dielectric to 0.05 volts. Calculate the ion current in E.S.U. and in amperes. If each ion carries 4.77×10^{-10} E.S.U. of charge, how many ions pass per second through the gauze?

27. An electrically charged cloud 2000 m on a side approached to within 100 m from the earth's surface when it discharged to a copper lightning conductor of resistance 3×10^{-2} ohms melting 400 grams of copper in the 1.55×10^{-3} seconds of the current flow and discharge. The latent heat of fusion of copper is 43 calories per gram and the copper was raised from 0° C to its melting point 1100° C with a specific heat 0.08. Calculate:

(*a*) The capacity of the cloud earth system in cms and in farads from the dimensions given, assuming the cloud and earth to form a parallel plate condenser.

(*b*) The current which flowed for 1.55×10^{-3} seconds and the quantity of electricity transferred.

(*c*) The potential between the cloud and the earth, and from this the sparking potential of a lightning discharge in volts/cm.

28. A condenser such as is used in telephone circuits having a capacity of 2 microfarads was charged to 1000 volts. It was then connected to a condenser of 4 microfarads capacity which was at 0 potential. Calculate the energy first in the charged condenser and then in the system of the two condensers. How much loss of energy was there, and how much were the connections raised in temperature if they weighted 0.05 gram and had a specific heat 0.08 calorie?

PROBLEMS BASED ON CHAPTERS XVII, XVIII AND XIX

1. The two vertical wires A and B each one meter long were suspended as shown in Fig. P23 with their ends dipping in two troughs of mercury TT. They were 1 cm. apart. A current of 22 amperes flowed through the two wires when they were in parallel (key K closed, switch S at P), and 20 amperes flowed through each when

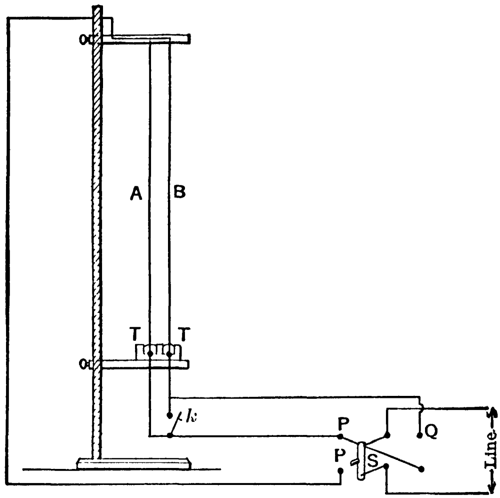


FIG. P23.

they were connected in series (switch S at Q, key K open). Taking g = 980 dynes, calculate:

(a) The force in grams between the conductors and the direction of motion when in parallel.

(b) The force in grams between the conductors and the direction of motion when in series.

2. A and B, Fig. P24, represent 2 exactly equal square coils each 100 cm on a side and separated by 5 cms, having their planes parallel and having the same axis perpendicular to their planes. A current of 10 amperes flows through them in exactly the same sense. Assuming that with the dimensions given the wires can be treated

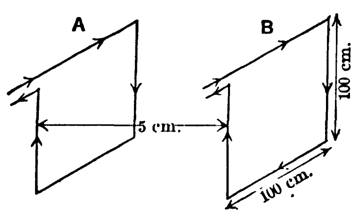


FIG. P24.

as infinitely long conductors, calculate the force between the coils in magnitude and direction, assuming that only the parallel wires act on each other. (From symmetry it is obvious that calculation need be made for one element of the coils only and multiplied by the appropriate constant.)

3. Given an infinitely long wire i carrying 20 amperes of current upward (see Fig. P25). It lies in the plane of a rectangular area $ABCD$ whose side AB is 2 cm from the wire i and whose side CD is 12 cm from i. The height of the rectangle is 30 cm. Calculate the flux in the area $ABCD$ due to the current in i. (*Note*: The flux $ABCD$ varies along AD. Thus φdA must be integrated over appropriate limits, where φ is the flux in a small vertical element of area dA, and φ varies with the distance from the wire.) What work in ergs would be done by moving a conductor 30 cm long carrying 20 amperes from AB to CD?

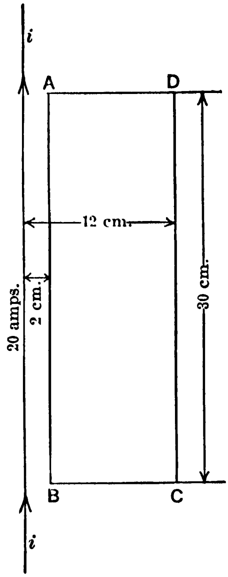


FIG. P25.

4. In the Barlow's wheel experiment there was a uniform magnetic field of 500 gauss over the whole radius of the wheel. The radius of the wheel was 10 cm, and 10 amperes were passed through it. Calculate the force on the wheel and the torque which was exerted.

5. For certain e/m experiments a solenoid 50 cm diameter and 400 cm long was used. It was covered with a single layer of cotton-covered No. 18 copper wire having 7 turns per cm and carried a current of 2 amperes. What was the field inside the coil and what was the total flux φ?

6. The radius of the coil in the demonstration experiment of the oscillating vertical coil (see Fig. P26) dipping into mercury is 2 cms. The turns of wire are 1 mm. apart. The force between 2 parallel wires with currents i_a and i_b is $\frac{2i_a i_b}{r} l$, where l is the length of one of the wires. If each cm of wire weighs 0.01 gram taking $g = 980$ dynes how much current must be run through the coil to make it oscillate against gravity? (While the contact with the mercury could for the spring extended by gravity at equilibrium be broken theoretically by a magnetic force of contraction of infinitely small amount, in practice the amplitude of oscillation to start the coil oscillating and to break contact is more nearly several mm. This requires that the force of gravity be practically overcome, that is, that each turn be just able to life the next turn against gravity and so permit the coil to contract by a sufficiently great amount.)

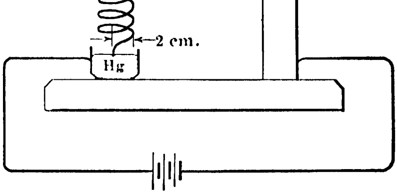


FIG. P26.

7. A straight bar of wrought iron 20 cm long whose area is 5 cm² is placed in a uniform coil where a field of 100 gauss exists. If the permeability of the iron for this field is 2000,

(*a*) Find the total flux through the bar.

(*b*) Assuming the magnetization concentrated at the end surface find I the intensity of magnetization.

(*c*) Find the pole strength of the temporary magnet.
(*d*) Find the susceptibility of the iron for this field.

8. A coil of wire 1 meter long with 1000 turns of wire carries a current of 2 amperes. A bar of iron of area 1 cm^2 and length 10 cm is placed in its center. If the permeability of the iron is 5000, calculate:

(*a*) The flux through the volume to be occupied by the iron in its absence.
(*b*) The total flux through the iron bar.
(*c*) Assuming the magnetism concentrated at the end surfaces of the iron, calculate the intensity of magnetization I.
(*d*) The pole strength of the temporary magnet.
(*e*) The susceptibility of the iron at this field strength.

9. A cathode ray beam or stream of electrons e of velocity v (Fig. P27), charge e, and mass m constitutes an electrical current flowing in the opposite sense to the

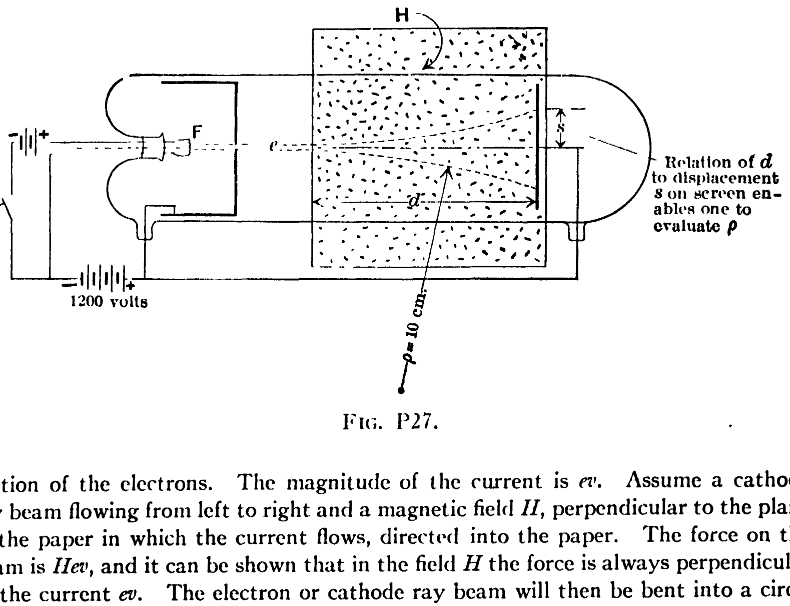


Fig. P27.

motion of the electrons. The magnitude of the current is ev. Assume a cathode ray beam flowing from left to right and a magnetic field H, perpendicular to the plane of the paper in which the current flows, directed into the paper. The force on the beam is Hev, and it can be shown that in the field H the force is always perpendicular to the current ev. The electron or cathode ray beam will then be bent into a circle of radius ρ. Since for circular motion the centripetal force is $\dfrac{mv^2}{\rho}$ we can write $Hev = \dfrac{mv^2}{\rho}$. If the electrons from a filament F have fallen freely through a potential of V volts before they enter H, they have a velocity v determined by $eV = \frac{1}{2}mv^2$. Given $V = 1200$ volts, $H = 11.64$ gauss, $\rho = 10$ cm, calculate:

(*a*) The sense in which the beam will be curved (up or down in figure).

(*b*) The value of v in terms of $\dfrac{e}{m}$.

(*c*) Put the value of v into the equation for ρ and solve for $\dfrac{e}{m}$.

(*d*) Then solve for v in cm/sec.
(*e*) Given $e = 4.77 \times 10^{-10}$ E.S.U. calculate m the mass of the electron.

The method outlined above is one in constant use in the study of electrical phenomena and is used to evaluate m the mass of the electron.

10. A beam of electrons constitutes a *negative* current of electricity given by e, where e is the charge and v the velocity of the electrons. In a magnetic field H the force is Hev and is always at right angles to the beam so that the electrons describe a circular path such that Hev equals the centripetal force mv^2/ρ, where m is the *mass* of the electron and ρ is the radius of curvature of the path. Given an *electron beam going from right to left in the plane of the paper and a field H = 15 gauss perpendicular* to the paper and *into it*, with

$$e = 4.8 \times 10^{-10} \text{ E.S.U.}$$
$$v = 5 \times 10^{9} \text{ cm/sec.}$$
$$H = 15 \text{ gauss.}$$
$$m = 9 \times 10^{-28} \text{ grams.}$$

Calculate ρ in cm and illustrate by diagram how the path will be curved.

11. An electromagnet of the design shown in Fig. P28 is to be built giving 4000 gauss across a gap 2 cm long. The area of cross section of the iron is to be 100 cm². The long side is 60 cm long, the short side is 40 cm long. There are to be two coils one on each of the 40 cm sides. The coils will not stand more than 4 amperes without overheating. The loss of flux in the gap is 25 per cent, and μ as a function of B is shown from the following data:

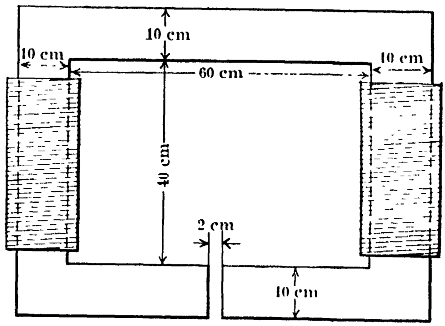


FIG. P28.

B	3506	4000	4434	4980	5340	5801
μ	3100	3200	3110	3000	2900	2750

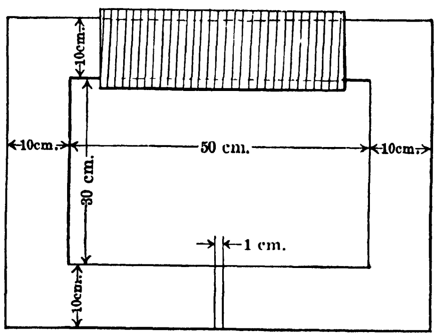


FIG. P29.

Calculate the number of turns of wire required.

12. Given the electromagnet on a wrought-iron frame shown in Fig. P29. The core has an area of 100 cm². The thickness of the iron is 10 cm. The inner length of the iron is 50 cm. The depth inside the frame is 30 cm. The air gap is 1 cm. Leakage is 20 per cent. The values of μ for this iron relative to B are given as:

B	3506	4000	4440	5000	5340	5801
μ	3300	3200	3110	3000	2900	2750

If there is to be a single coil which can carry only 3 amperes and the field in the gap is to be equal 4000 gauss, what is:

(*a*) The reluctance of the circuit.
(*b*) The magneto-motive force, and
(*c*) The number of turns needed in this case.

13. An iron bar 40 cm long and 10 cm² cross section is bent into the shape of a horseshoe for the purpose of making an electromagnet which shall have a pull of 50 kg upon its armature (a bar 12 cm long and 10 cm² cross section), when it is 1 cm away from the poles. Find the ampere turns required if the leakage is 20 per cent. The values of B and μ for the iron are as follows:

B........	5000	9000	10,000	11,000	12,000	13,000	14,000	15,000
μ.........	3000	2250	2000	1700	1400	1100	825	525

14. In a transformer of the type depicted in Fig. P30, where the secondary coil II is wound on a continuous laminated square iron core which also carries the primary coil I, there is a magnetic shunt provided by adding a second iron pathway indicated by ABA in the figure. When the bar B is absent, all the flux in I goes through II. When B is in place between AA the flux is in part diverted through the path ABA and thus reduces the E.M.F. obtained in II. Given the iron from between I and II as 50 cm long and 30 cm wide inside the frame with an area of 10 cm $\times$ 10 cm cross section, calculate the value of the reluctance of the shunt required to reduce the flux φ_0 in the absence of the shunt (i.e., of B), to $\frac{1}{2}$ the value when B is in place. From the value found, design or calculate the dimensions of such a shunt. For the iron used, the problem may first be solved for the case when μ is constant and 2000 in the branch II and in ABA whether B is in place or not. It may more correctly be solved using $\mu = 2000$ when B is absent, and $\mu = 3000$ in II when B is in place, while the value of μ for B is 3000.

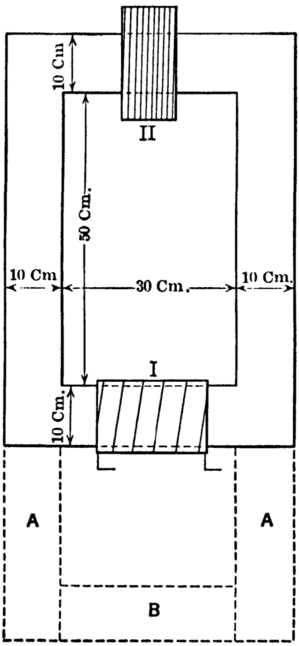


FIG. P30.

15. In steel mills they employ powerful electromagnets for lifting pieces of steel and transferring them to freight cars. Such a magnet has a length of 400 cm, and an area of cross section of 1600 cm². It is to lift steel ingots 40 $\times$ 40 cm weighing 1600 kg. When across the magnet the flux runs through 100 cm of the ingot. Calculate the value of B required to lift this mass. For this value of B, μ is 500. Calculate the ampere turns required to produce the flux, neglecting leakage, and if the allowable current is 80 amperes, calculate the number of turns required.

16. A horseshoe magnet is to be designed to lift a load of 40 kg. on a keeper bar the dimensions of which are given as follows, in a magnet of the type depicted in

Fig. P31. The length of the horseshoe part is 45 cm. The length of the keeper is 15 cm. The horseshoe has a cross section of 5 × 5 cm. The keeper has a cross section of 5 × 4 cm. The air gaps A, A, are each 0.01 cm long. The loss of flux is 10 per cent. The coil C cannot take more than 1 ampere without overheating. The values of B and μ for the iron used are as follows:

B......	3506	4000	4440	4935	5340	5801
μ.......	3300	3200	3110	3010	2900	2750

Then calculate the following data assuming the flux through the keeper is confined to an area 5 × 5 cm at the ends of the magnet.

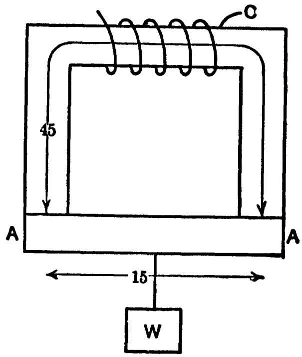


FIG. P31.

(a) The induction B going through the keeper to hold the weight. (Remember that the weight of 40 kg is supported at two poles A and A, as well as that the area of cross section of the keeper is less than the area of the entering flux from the horsehoe.)
(b) The flux ϕ through the keeper.
(c) The values of μ for the horseshoe and keeper from B for each.
(d) The reluctances of the horseshoe, the keeper, and the air gaps.
(e) The ampere turns required to lift the weight.
(f) The number of turns required at 1.0 ampere through the coil.

17. Given the circuit depicted in Fig. P32, with the data required for calculating Z_1, and Z_2 given as follows:

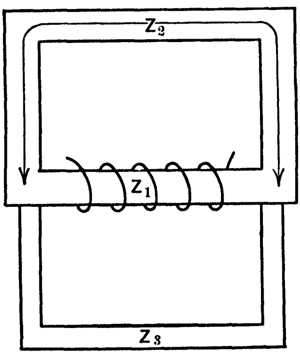


FIG. P32.

Length of conductor $Z_1 = 20$ cm.
Area 5 × 7 cm, $\mu = 300$.
Length of conductor $Z_2 = 80$ cm.
Area 5 × 5 cm, $\mu = 350$.

A shunt Z_3 is to be constructed so that ϕ_3 through it shall be 1/10 of ϕ_2 the flux through Z_2. The coil producing the flux through the circuit is about Z_1 and gives a M.M.F. having the values of μ above for Z_1 and Z_2. About Z_3 the following information is known. The area of cross section of Z_3 must be 2 × 2 cm, and with the fact that $\phi_3 = (1/11)\phi_1$ and that its area is 4 cm² relative to the 35 cm² for Z_1 the value of B and hence of μ is found to be 500. It is then required first to find the value of Z_3 and then its length so that the diversion of flux required can be made. To solve this problem you are given that M.M.F. $= \phi_1(Z_1 + Z')$ where Z' is the combined reluctance of Z_2 and Z_3 in parallel, and that M.M.F. $= \phi_3 Z_3$, while $\dfrac{\phi_3}{\phi_1} = 1/11$. With these data solve for Z_3 and l_3 the length of Z_3.

18. (*Note:* In solving this problem do not multiply out or divide by π as the values of π appearing should cancel.) A transformer has a continuous iron core of 40π cm length, $5 \times 4 = 20$ cm² area cross section, and a value of $\mu = 2000$. A *primary* coil

has 50 turns and di amperes is the change of current in dt seconds, as will be seen below. The *secondary* coil has 2×10^5 turns. The transformer is fed by a 60-cycle per second alternating current ($N = 60$), i.e., one with $T = \frac{1}{60}$. For an alternating current $i = i_0 \sin\left(\frac{2\pi t}{T}\right)$ whence $\frac{di}{dt} = \frac{2\pi}{T} i_0 \cos\left(\frac{2\pi}{T} t\right)$ which is a maximum when cos $\frac{2\pi t}{T} = 0$, the time at which sparking occurs. Hence $\frac{di}{dt} = \frac{2\pi}{T} i_0$. With these data calculate:

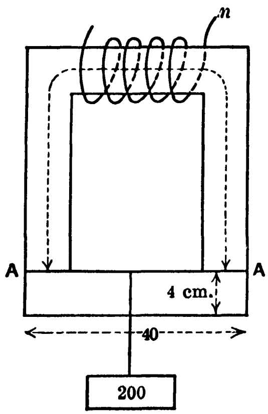


FIG. P32.1.

(*a*) The reluctance Z of the transformer.

(*b*) $\frac{d\phi}{dt}$ in the iron in terms of i_0.

(*c*) E in the secondary expressed as volts in terms of i_0.

(*d*) The value of i_0 in the primary necessary to give a potential of $10{,}000\pi$ volts in the secondary (i.e., that to give a 1 cm spark in air).

19. A horseshoe magnet is to be designed to lift a load of 200 kg on a keeper bar the dimensions of which are given below, the magnet being as shown in Fig. P32.1. The area of contact between keeper and magnet is 7×7 cm, which is the area of cross section of the horseshoe magnet, for the air gaps A, A are only 0.015 cm long. The keeper is 40 cm long and has an area of cross section 7×4 cm. The length of the horseshoe is 150 cm. The leakage at the gaps is 20 per cent of the flux in the horseshoe. The values of μ and B for the iron used are given below. Calculate the ampere turns required.

B	7000	8000	9000	10,000	11,000	12,000	13,000
μ.	2200	2000	1800	1,600	1,300	1,100	950

PROBLEMS BASED ON CHAPTERS XX, XXI AND XXII

1. A Faraday disc 10 cm radius is rotated at 1200 R.P.M. and develops 2 volts E.M.F. What is the strength of the field, assuming it to be uniform over the radius?

2. A secondary coil of 100 turns is wound about a primary coil of 6000 turns, area 5 cm² and length 20 cm. An unknown current is suddenly broken in the primary. The ballistic galvanometer with resistance of 100 ohms including the resistance of the secondary coil indicated that 7.536×10^{-6} coulombs had passed through the secondary during the time of breaking. Calculate: (*a*) The current flowing in the primary; and (*b*) the field existing in the primary when the current flowed.

3. A coil of 10 turns, with a mean diameter of 1 cm, is connected to a ballistic galvanometer, the resistance of the complete circuit being 5 ohms. The coil is placed between the poles of an electromagnet initially with its plane normal to the field and then quickly turned through 180°. It is noted that 3.1416×10^{-6} coulombs of electricity pass through the galvanometer. Determine the field between the poles of the magnet.

4. A solenoid, 50 cm long and 1 cm internal diameter, has a total of 1000 turns of wire. A secondary of 100 turns is would around the middle of the solenoid and connected to a ballistic galvanometer. The resistance of the secondary circuit including

the galvanometer is 10 ohms. When a current of 5 amperes is made or broken in the solenoid what quantity of electricity is set in motion in the secondary?

5. Suppose in the last problem an iron rod just filled the solenoid and that the permeability of the iron for the field was 300. What quantity of electricity would now be set in motion in the secondary when 5 amperes were made, or broken, in the solenoid?

6. A coil of wire 20 cm long with 1000 turns carries a current of $\frac{1}{2}$ ampere. A bar of iron, area of cross section 2 cm^2 and length 20 cm is placed in its center. If the permeability of iron is 2000, calculate:

(*a*) The flux through the volume to be occupied by the iron before it is introduced.

(*b*) The flux through the iron bar when in the field.

7. The earth inductor shown in the lecture demonstration has 100 turns and a radius of 10 cm. Its resistance together with the galvanometer was 50 ohms. If 3.14×10^{-6} coulombs of electricity passed in $\frac{1}{2}$ a revolution, calculate H_T, the total intensity of the earth's field. If the dip was 75°, calculate the horizontal intensity of the earth's field.

8. The earth inductor shown in the lecture has 50 turns and a radius of 20 cm. Its resistance with that of the ballistic galvanometer was 100 ohms. If 1.57×10^{-6} coulombs were observed to flow on $\frac{1}{2}$ a rotation, calculate the total intensity H_T of the earth's field. If the dip was 70° what was the horizontal component?

9. In the demonstration experiment a secondary coil of 20 turns is wound about a primary coil of 200 turns having an area of 2 cm^2 and length 40 cm. A current of 0.5 ampere is suddenly broken in the primary.

(*a*) How many coulombs passed through a galvanometer whose resistance with that of the coil was 100 ohms?

(*b*) Had a piece of iron $\mu = 2000$, and 20 cm long with an area of 1 cm^2 been in the coil what would the quantity in coulombs have been?

10. Given the total intensity of the earth's field as about 0.4 gauss in a region It is to be studied by means of a ballistic galvanometer and an earth inductor radius 20 cm and 100 turns. The deflection of the galvanometer is read by telescope and scale, the scale being one marked in mm is 1 meter distant from the galvanometer. The galvanometer and coil have 100 ohms resistance. Calculate the figure of merit of the galvanometer to be used in order to measure H_T to 1 per cent if only $\frac{1}{2}$ revolution is to be used, neglecting the damping factor and assuming that the reading can be estimated to mm only.

11. Magnetic fields in solenoids are frequently measured by " flip coils " which are merely small analogues of the earth inductor worked by a spring and ratchet arrangement so that they always make $\frac{1}{2}$ revolution in the same time and can be conveniently automatically released. Such a coil has a radius of 0.5 cm and 500 turns of wire with a resistance of 1 ohm, while the galvanometer has a resistance of 19 ohms. Give the specifications of the sensitivity and figure of merit of a galvanometer to be used at 1 meter with a scale divided in mm in order to measure fields of 50 gauss to 1 per cent. The period of the galvanometer is 20 seconds.

12. In an experiment, a wire is moved along rails across a field from a powerful electromagnet, the rails were 10 cm apart. When the wire was moved 2 cm in 0.2 second a deflection of 20 cm on a scale at 1 meter distance was obtained when

$K = 10^{-4}$, $T = 2$ seconds, ρ was negligible, and R was 20 ohms. What was the number of lines cut and what was the average value of H in the region cut?

13. Five hundred turns of wire are wound on a square frame, the average length of each turn being 160 cm. The frame makes 300 R.P.M. about a horizontal axis, through its center and parallel to an edge. There is a uniform horizontal magnetic field of 10 gauss at right angles to the axis of rotation. The initial position is taken when the plane of the frame is normal to the field. The wire on the frame is connected through ring connectors with an external circuit, the resistance of the complete circuit being 10 ohms.

(a) Determine the quantity of electricity set in motion for half revolutions, starting from the initial position.

(b) What is the average current during this half revolution? Will this be the average current for an indefinite time the rate of rotation being constant at 300 R.P.M.?

(c) Determine the minimum E.M.F. and the maximum E.M.F. in volts, and the positions of the frame at the instant these values are obtained.

(d) Determine the E.M.F. in volts at the instant the frame makes an angle of 30° with its initial position.

(e) Find the average E.M.F. and the virtual E.M.F.

14. A motor has a resistance of 1 ohm and operates on a 100-volt circuit. When operating with a current of 10 amperes, what is the back or counter E.M.F.? What is the electrical efficiency of the motor?

15. A motor runs on an applied potential of 110 volts. It draws 4 amperes of current and its internal resistance is 2 ohms. Calculate:

(a) The back E.M.F.

(b) The power consumed in useful work.

(c) The efficiency, neglecting friction.

(d) If it makes $\dfrac{6000}{\pi}$ R.P.M. what is the force on a pulley 5 cm radius in Kg?

(e) If it were run as a dynamo at the above speed and gave a current of 2 amperes, what P.D. would there be at its terminals?

16. A motor is run on 120 volts D.C. It draws 5 amperes and its internal resistance is 1 ohm. Calculate:

(a) The back E.M.F.

(b) The power consumption in useful work if 4.8 ampere go to this work and 0.2 ampere to hysteresis and other losses.

(c) Given a current at the operating voltage of 0.2 amperes which goes to frictional losses, etc., out of the 5 amperes used. Calculate the efficiency of the motor under the given conditions.

(d) If it makes $\dfrac{5000}{\pi}$ R.P.M. what is the forçe on a pulley of 6 cm radius in kg?

(e) If it were run as a dynamo at the above speed and gave a current of 4 amperes what P.D. would there be at its terminals, neglecting eddy current losses?

17. The secondary of an induction coil with iron removed is to be used as an inductance. It has 10^5 turns a length of 31.4 cm, and an area of 500 cm². If $i = 0.1$ ampere, calculate:

(*a*) The field H.

(*b*) The total flux ϕ which threads it.

(*c*) If the current be broken in 1 sec, the E.M.F. in the 10^5 turns in E.M.U. and volts.

(*d*) The self-induction in cm and henrys.

18. A coil of unknown number of turns n when a current of 0.1 ampere is suddenly broken in it due to its self-induction delivers a quantity of 0.2 coulomb through a galvanometer of resistance of 1000 ohms. The coil is 31.4 cm long, has no iron in it and has an average area of cross section of 500 cm.^{2} Remembering that when the field H due to the current collapses the flux cuts the n turns of the wire to generate the quantity 0.2 coulomb, and taking $\pi = 3.14$, calculate n the number of turns of wire in the coil.

PROBLEMS BASED ON CHAPTERS XXIII, XXIV, XXV AND XXVI

1. The coefficient of self-induction L is defined by the equation $E = L\dfrac{di}{dt}$. For a coil of n turns of length l, area A, with permeability μ and a current i amperes in it the change of flux is $\dfrac{4\pi}{10}\dfrac{niA\mu}{l}$. When the current ceases, $\dfrac{d\phi}{dt} = \dfrac{4\pi nA\mu i}{l\tau}$ if the change occurs in τ seconds. The E.M.F. induced in the n turns of the coil by this change due to cutting of its own wires will be

$$E = \frac{4\pi}{10}\frac{n^2A\mu}{l}\frac{i}{\tau}. \quad \text{If } \frac{i}{\tau} = \frac{di}{dt}$$

this expression enables one to calculate L for the coil. Given a coil 20 cm long with 4000 turns and an area of 10 cm^2. In one experiment it has no core, in the other it has a core for which the average μ is 4000. Calculate L the coefficient of self-induction (*a*) when iron is absent, (*b*) when it is present, (*c*) in how short a time must the current be broken to give an E.M.F. of 2 volts?

2. It is required to have a ballistic galvanometer that will throw a spot of light 200 mm on a scale 1 meter distant when a quantity of 10^{-8} coulombs passes through it. The damping on successive swings to the right should not be more than in the ratio of 21/20, and the period must exceed 2 seconds for a complete swing. Calculate the figure of merit and sensitivity of the galvanometer in megohms so that in going over the catalogues you would be able to order the proper galvanometer.

3. A parallel plate condenser of area 3140 cm^2 and a plate distance of 0.02 cm achieved by a thin plate of mica of dielectric constant $D = 6$ is charged to a potential of 100 volts. On discharge through a ballistic galvanometer of figure of merit $k = 4.175 \times 10^{-8}$, with a period of 8 seconds, this gives a deflection of 15.7 cm corrected for damping.

(*a*) Calculate C the capacity in E.S.U.

(*b*) Calculate Q the charge on the condenser.

(*c*) Calculate C the capacity in absolute E.M.U.

(*d*) From the ratio of the values of $\dfrac{C \text{ in E.S.U.}}{C \text{ in E.M.U.}}$ obtain the ratio of the units of capacity in the E.M. and E.S. system $\left(\text{i.e., } \dfrac{C \text{ e.m.u.}}{C \text{ e.s.u.}}\right)$. Take $\pi = 3.14$ and

note that this method is one of the methods of obtaining the ratio of the E.S.U. and E.M.U.

4. Given a circular iron ring of mean length 100 cm and area of cross section of the iron 10 cm², having a coil of 1000 turns and $\mu = 4000$. Calculate the self-induction of the coil. If the coil be closed through a ballistic galvanometer so that the resistance is 10 ohms what quantity of electricity will flow on breaking a current of 2 amperes?

5. An E.M.F. of 30 volts is applied to a coil of 3 ohms resistance and 0.5 henry inductance.

(*a*) What is the time constant of the coil?

(*b*) What is the current in the coil 0.01 sec; 0.1 sec; 1 sec; 10 sec after the E.M.F. is applied?

6. A coil, without an iron core, used on the lecture table, has a resistance of 1.6 ohms and an inductance of 33 millihenrys. Determine the "reactance" and the "impedance" for an alternating current of 60 cycles. If a 60-cycle alternating E.M.F. of 110 volts is applied to the coil what current will be obtained? If the core is introduced the self-induction is 3 henrys. What are the values of the quantities in the presence of the core?

7. The resistance of the carbon filament lamp used for the demonstration of the effect of capacity on a circuit may be assumed constant in what follows and equal to 100 ohms. One hundred and ten volts alternating potential of 60 cycles are applied across the lamp and condenser in series. When the capacity is 10 microfarads and 1 microfarad, calculate:

(*a*) The current through the lamp in each case.

(*b*) The angle of phase lag in each case in degrees.

8. A voltmeter-ammeter reading was taken on an alternating E.M.F. of 60 cycles applied to a circuit with self-induction, resistance and capacity. The voltmeter read 100 volts; the ammeter read 5 amperes. A wattmeter in the same circuit read 200 watts. What was the power factor and the angle of phase lag in the circuit? If the capacity was 5 microfarads, the resistance was 100 ohms, calculate the value of L the coefficient of self-induction of the system.

9. Given an alternating P.D. of 158 volts amplitude, applied to a circuit having a coefficient of self-induction $L = 2$ henrys, a frequency of 60 cycles and a resistance of 240π ohms. Calculate:

(*a*) The virtual voltage.

(*b*) The angle of phase lag.

(*c*) The virtual current and impedance.

(*d*) The power factor.

10. A coil has a coefficient of self-induction of 2 henrys. If it has 10 ohms resistance and is suddenly attached to 100 volts, what is the current .01, .1, 1.0, and 10 seconds after the switch is closed? Were a current of 1 ampere suddenly interrupted and the coil connected to a galvanometer, what would the current read .01, .1, 1.0, and 10 seconds after the circuit was broken?

11. A ballistic galvanometer to be used with an earth inductor giving quantities of the order of 10^{-7} coulombs is needed. The periods of such instruments are of the order of 4 seconds, and the damping is such that on two successive swings to the same

side the amplitudes are in the ratio of 51/50. Calculate the figure of merit and sens tivity in megohms of a galvanometer to be used for the above purpose so that it wi give 100 mm deflection on a scale 1 m distant for 10^{-7} coulombs.

12. In a certain experiment where high resistances are needed these are made b drawing an india ink line on some suitable backing making contact with two met leads. Such a resistance was made and its value was measured using the time c discharge of a condenser of 1 microfarad. When the condenser was charged to suitable potential and discharged at once the throw was 20 cm. After short-ci cuiting the condenser for 20 seconds through the india ink resistance for the sam initial charge the throw was 15 cm on the scale. Calculate the resistance R in ohm If 10^6 ohms are called 1 megohm calculate the resistance in megohms.

13. Given a circuit with 2 henrys self-induction. How long after closing a swit will it take with a resistance of 20 ohms before the current has risen to 50, 90 and 9 per cent of its full value?

14. Given a P.D. of 141.6 volts amplitude applied to a circuit having a coefficien of self-induction L of 1 henry, a frequency of 60 cycles, a capacity of 0.5 microfarad and a resistance of 200 ohms. Calculate:

(*a*) The virtual voltage.
(*b*) The impedance and virtual current.
(*c*) The angle of phase lag or advance, stating which.
(*d*) The power factor.

15. A given A.C. circuit has the following constants, a capacity C of 10 microfara a resistance R of 82 ohms, a period factor p of 360, an unknown self-induction L An ammeter-voltmeter reading gave 120 volts and 1.465 amperes. A wattmete reading of the same circuit gave 120 watts. Calculate:

(*a*) The power factor.
(*b*) The angle of phase lag and its tangent.
(*c*) From the tangent of the angle of phase lag, *assuming* that there is a *la* not an advance (i.e., ϕ is negative) calculate L the unknown self-induction.

16. Given an A.C. circuit having a resistance $R = 82$ ohms, a capacity C of 1 microfarads, a period factor p of 360, and self-induction L of 1 henry. Calculat the following data:

(*a*) The impedance and reactance.
(*b*) The virtual current if the impressed virtual voltage was 120 volts.
(*c*) The angle of phase lag.
(*d*) The power factor.

17. Radioactive substances decay according to a law $N_t = N_o e^{-\lambda t}$, where t i the time in seconds, λ is the constant characteristic of a given change, N_t is the nur ber of atoms at a time t and N_o is the initial number of atoms at $t = 0$. The rate c growth of the resultant substance is $N_t = N_\infty(1 - e^{-\lambda t})$, where N_∞ is the ultimat number of particles after all the parent substance has transformed. λ for radium i 1100 (years^{-1}). If radium changes at this rate to emanation and emanation unde goes change to radium A, B, C and D at rates far greater than the rate of change c radium, the amount of radium transformed in a year will have undergone practicall complete change to D in this time. In the changes involved from Ra to RaD ho many α particles are given off? Starting with 1 gram of radium having $\frac{6.06 \times 10^{\cdot}}{226}$

atoms how many atoms undergo change in the course of a year? If each of these gives the same number of α rays, compute how many mm^3 of *He* at N.T.P. (number of atoms in 1 cm^3 = 2.705×10^{19} at N.T.P.) are formed from a gram of *Ra* in a year? Check this value against that computed from the fact that 1 gram of radium gives out 1.36×10^{11} α particles per second, in its change to *RaD*. The discrepancy shows the range of errors in the measurement of the life of radium and in the number of α particles per second.

18. Scintillation counts show that 1 gram of radium produces 1.36×10^{11} α particles per second. Rutherford let the fraction of α particles from 10 mg of radium that entered a cone of 1 cm in diameter at 10 cm distance fall on a Faraday cylinder in vacuum connected to an electrometer, the capacity of the system being 100 cm. It was observed that the electrometer indicated a charge of 2.43 volts in 16 minutes and 40 seconds. Calculate the charge on an α particle. If the electron has a charge of 4.77×10^{-10} E.S.U., what multiple of this charge does the α particle carry?

19. In the thorium series of disintegrations, given the following sequence of changes:
Th α Meso Th1 β Meso Th2 β Radio Th α ThX α Th Em α Th A α Th B β Th C β Th C′ α Th D. If thorium falls in Group IV (the lead group) and in the 7th period having an atomic weight 232, deduce the chemical behavior of the transformation products down the chain and deduce the atomic weight of Th D as well as its chemical behavior.

20. Given the atomic number of rubidium as 37. Construct a rough diagram of the atom showing the disposition of electrons in levels, and giving approximate dimensions of the various quantities entering in. Do the same for iodine, I, atomic number 53. On this basis interpret the electrochemical behavior of Rb and I.

21. Calculate the wave lengths of the following radiations from the potential through which an electron must fall to generate them. Given $h = 6.56 \times 10^{-27}$ erg sec, and the electronic charge as 4.77×10^{-10} E.S.U.

2537 Å mercury line	4.9 volts
K_α line of carbon	288 volts
K_α line of tungsten	57,200 volts
K_α line of uranium	99,500 volts
Millikan's hard x-rays	750,000 volts
γ rays RaC very hard	1.79×10^6 volts

TABLE OF UNITS AND DIMENSIONS

Entity Fundamental or Derived	Defined by the Following Equation	Dimensions	Electromagnetic Unit
CURRENT FUNDAMENTAL: Comes from Electro-Magnetic System	$f = \frac{ids}{r^2}$ ds = normal to r r = distance i = current f = force on unit magnet pole	In E.M System, $i = M^{1/2}L^{1/2}T^{-1}$ In E S. System, $i = M^{1/2}L^{3/2}T^{-2}$	$i = 1$ when $ds = 1$ cm normal to r $r = 1$ cm $f = 1$ dyne on unit pole Also $f = ilH$ $i = 1$ when $l = 1$ cm $f = 1$ dyne $H = 1$ gauss Large
QUANTITY FUNDAMENTAL Comes from Electrostatic System	$f = \frac{qq'}{Dr^2}$ q and q' = quantities f = force between them r = distance D = dielectric const	In E M. System, $q = M^{1/2}L^{1/2}$ In E S System, $q = M^{1/2}L^{3/2}T^{-1}$	$q = it$ $q = 1$ when $i = 1$ E.M.U. $t = 1$ second Large
POTENTIAL FUNDAMENTAL Comes from work in either system	$W = q \times$ P D or W = P D it W = Work q = Quantity i = Current t = Time P.D = Potential	In E M System, P D $= M^{1/2}L^{3/2}T^{-2}$ In E S System P.D. $= M^{1/2}L^{1/2}T^{-1}$	P D = 1 when $q = 1$ E M U. $W = 1$ erg P D $= \frac{d\phi}{dt}$, $\frac{d\phi}{dt}$ = no lines force cut by conductor per sec P D. = 1 when $\frac{d\phi}{dt} = 1$ Small
RESISTANCE DERIVED: From E M. System A constant of the circuit shape, and material	$R = \frac{\text{P D}}{i}$ R = Resistance P D. = Potential i = Current	In E M System, $R = LT^{-1}$ a velocity In E.S System $R = L^{-1}T$ a reciprocal of velocity	R = P D , i $R = 1$ when P D = 1 E M U $i = 1$ E M U Very small
CAPACITY DERIVED: From E S. System A constant of the shape and material surrounding circuit	$C = \frac{q}{\text{P.D}}$ C = Capacity q = Quantity P.D. = Potential	In E M. System, $C = L^{-1}T^2$ In E S. System, $C = L$ Unit is the cm	$C = \frac{q}{\text{P D}}$ $C = 1$ when $q = 1$ E M U P D. = 1 E.M.U Very large
SELF-INDUCTION DERIVED: Defined from E M. system A constant of circuit depends on shape and magnetic media involved	P.D. $= L\frac{di}{dt}$ P.D. = Potential generated $\frac{di}{dt}$ = rate of change of current L = Coefficient of self-induction	In E.M System, $L = L$ Unit is the cm In E.S System, $L = L^{-1}T^2$	P.D. $= L\frac{di}{dt}$ $L = 1$ when P D = 1 E.M.U. $\frac{di}{dt} = 1$ E.M.U./sec. Very small the cm

Electrostatic Unit	Practical Unit	Relation Between E.M.U. and E.S U.	Relation Between E M U and Practical	Relation Between E.S.U and Practical
$i=\frac{q}{t}$ $i=1$ when $q=1$ E.S.U. $t=1$ sec. Small	Ampere defined as 0.1 of E.M.U	1 E.M.U. $=3\times10^{10}$ E S U	1 E.M.U. = 10 amperes	1 Ampere $=3\times10^{9}$ E.S.U.
$f=\frac{qq'}{Dr^2}$ $q=1$ when $q=q'$ $r=1$ cm $D=1$ $f=1$ dyne Small	Coulomb defined as 1 ampere for 1 sec or as 0 1 E M U of q	1 E M U. $=3\times10^{10}$ E S U	1 E M U = 10 Coulombs	1 coulomb $=3\times10^{9}$ E S U.
$W=$ P D. $\times q$ P D = 1 when $W=1$ Erg. $q=1$ E S U Large	Volt defined as 10^8 abs E M U	1 E S U $=3\times10^{10}$ E M U	1 volt $=10^8$ E M U	1 E S U = 300 volts
$R=$ P D $/i$ $R=1$ when P D = 1 E S.U $i=1$ E S U Very large	Ohm defined as $\frac{\text{amperes}}{\text{volts}}=$ Ohms	1 E S U. $=9\times10^{20}$ E.M U	1 Ohm $=10^9$ E M U	1 E S U $=9\times10^{11}$ Ohms
$C=\frac{q}{\text{P D.}}$ $C=1$ when P.D. = 1 E S.U. $q=1$ E S.U. Very small the cm	Farad defined as $\frac{\text{coulombs}}{\text{volts}}=$ farads. Common unit the microfarad $=10^{-6}$ farads	1 E M U 9×10^{20} E.S.U.	1 E M.U $=10^9$ farads $=10^{15}$ microfarads	1 farad $=9\times10^{11}$ E S U. $=9\times10^{11}$ cm. 1 microfarad $=9\times10^{5}$ cm 1 micro microfarad = 0.9 cm
P.D. $=L\frac{di}{dt}$ $L=1$ when P.D. = 1 E S U. $\frac{di}{dt}=1$ E S.U./sec Very large	Henry defined as $\frac{\text{volts}}{\text{amperes/sec}}=$ henrys	1 E.S U. $=9\times10^{20}$ E M.U	1 henry $=10^9$ E M.U $=10^9$ cm 1 millihenry $=10^{-3}$ henry $=10^6$ cm 1 micro henry $=10^{-6}$ henry = 1000 cm	1 E S U $=9\times10^{11}$ henrys

AUTHOR INDEX

SUBJECT INDEX

E

M

N

T